Fall 2014

Effect of Grain, Region, and Human Influence on Higher Taxonomic Surrogacy

Frances Owen
Fort Hays State University, contact.francesowen@gmail.com

EFFECT OF GRAIN, REGION, AND HUMAN INFLUENCE

ON HIGHER TAXONOMIC SURROGACY

being

A Thesis Presented to the Graduate Faculty

of the Fort Hays State University in

Partial Fulfillment of the Requirements for

the Degree of Master of Science

by

Frances Owen

B.S., University of Dubuque

Date_____ Approved_____
Major Professor

Approved_____
Chair, Graduate Council

This Thesis for

The Master of Science Degree

by

Frances A Owen

has been approved

_____
Chair Supervisory Committee

_____
Supervisory Committee

_____
Supervisory Committee

_____
Supervisory Committee

_____
Chair, Department of Biological Sciences

PREFACE

This manuscript has been formatted in the style of the journal Conservation Biology.

Keywords: surrogacy, higher taxa, biodiversity, grain, ecoregion, human influence

ABSTRACT

Surrogacy is a common tool within conservation and can be useful when scientists lack detailed knowledge of a system. Higher taxonomic surrogacy is appealing because it can save time and money. However, this technique might vary in effectiveness depending on the taxonomic level, spatial grain, region, and impact by humans. In this thesis I addressed some of the common concerns with higher taxonomic surrogacy using Breeding Bird Atlas data from six states (Colorado, Florida, Michigan, New York, Pennsylvania, and Washington). I compared the coefficients (slopes) of my models rather than the R^2 values relied on by other higher taxonomic surrogacy studies. My results suggest taxonomic level, spatial grain, and region can affect higher taxonomic surrogacy. I did not detect a clear pattern between higher taxonomic surrogacy and human influence. I conclude that higher taxonomic surrogacy is a potentially useful tool for assessing biodiversity in an area, but care should be exercised when using this surrogacy technique to predict biodiversity.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor Dr. Rob Channell for his guidance as I worked through my thesis. I could not imagine a better person to have accompanied me on this journey. Besides my advisor, I would like to thank the rest of my committee Dr. Greg Farley, Dr. William Stark, and Dr. William Busby for their support and patience throughout the process.

Everyone in the Department of Biological Science was a huge help, either directly or indirectly. I would like to thank all of the graduate students in particular for their constant inspiration and for distracting me when I needed it.

I want to acknowledge the thousands of professional and citizen scientists responsible for contributing the data (especially the breeding bird atlas data) I used for this thesis. Without them, this project would not have been possible.

I thank all of my friends and family for their continued support especially my parents, David and Vickie Eggers, whose trust and encouragement greatly shaped the person I am today. Most of all I want to thank my husband for always believing in me when I wanted to go to graduate school and for keeping me sane after I started.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Earth is facing a biodiversity crisis. Extinction rates are currently 100 to 1000 times natural background rates and some scientists suggest we have entered a sixth mass extinction event (Pimm, 1995; Mora et al. 2011; Barnosky 2011). Biodiversity loss is the greatest issue facing conservation biologists today. Comprehensive biological inventories are great tools for understanding changes in biodiversity within an area; however these inventories can be expensive and time consuming because they typically require taxonomic experts to accurately identify organisms to the species level (Marshall et al. 2006, Jones, 2008). Approximately 1.2 million species have been described by taxonomists (Mora et al. 2011). By some estimates, this suggests more than 7.5 million species have not been described (Mora et al. 2011). Other global estimates range from 2 to 50 million total species (Stork 1993). One study estimated that it would take thousands of taxonomists over 1,000 years and over $350 billion U.S. dollars to describe the remaining unknown species (Mora et al. 2011). High extinction rates mean that species will become extinct without being described and named; this is a centinelan extinction (Wilson 1992). Therefore, we need economical methods to efficiently and accurately measure biodiversity as well as methods that could potentially monitor undescribed species richness. Some studies have suggested that identification to the species level might not be necessary for understanding the spatial distribution of species richness and have suggested using surrogates of species richness to assess biodiversity.

The most common types of surrogates include flagship species, umbrella species,

and indicator species (Andelman & Fagan 2000; Caro & O'Doherty, 1999). Flagship species are charismatic species, typically large mammal or bird species, used to garner public and financial support for conservation initiatives (Williams et al. 2000; Western 1987). They are sometimes incorrectly used as umbrella species (Caro & O'Doherty 1999). Umbrella species require large areas of habitat and by conserving habitat for the umbrella species you also benefit many other species that use the same habitat (Noss 1990; Caro 2003; Roberge & Angelstam 2004). These species are sometimes used in reserve planning, however, it can be difficult to assess if an umbrella species will be effective at protecting other species of concern (Rubinoff 2001; Tracy & Brussard 1994). An indicator species is an organism used as an indirect measure of a factor of interest, including biodiversity or environmental conditions (Landres et al. 1988; Caro & O'Doherty 1999). In the past, indicator species have been used primarily for assessing pollution, but recently indicators have been used to monitor population trends in other species (Landres et al. 1988).

Another surrogate technique, higher taxonomic surrogacy, could be a useful conservation tool. Higher taxonomic surrogacy uses higher taxa richness, such as genus or family richness, to predict species richness. This method is potentially better for monitoring biodiversity than the previously mentioned surrogates, because it more directly represents the species diversity in an area. Higher taxa such as genera and families might conserve more data about species ecology and evolution than do umbrella species or indicator species (Warwick et al. 1995). However, the data maintained in

higher taxonomic levels might not be ecologically meaningful (Stanislao et al. 2012). The rationale behind higher taxon surrogacy is that each taxonomic level is nested within a less diverse higher taxonomic level. Thus, there are fewer higher taxa than lower taxa, and using higher taxa to predict species richness could result in a reduction of time and effort spent identifying species.

Although higher taxonomic surrogacy has only recently gained attention in conservation, the method has been used within paleontology for some time (Allmon 1992). Paleontologists have used higher taxa richness to predict lower taxa richness and monitor prehistorical extinction events (Gaston & Williams 1993). There is a bias within the fossil record regarding the types of specimens that are preserved (Signor 1990). Organisms that are large, possess calcified parts, and species that are deposited in an appropriate environment, will have a better chance of fossilizing than would small soft-bodied organisms. Fossil formation requires a specific series of rare events, therefore it is more probable that at least one fossil specimen from a family will be preserved rather than every species within that family being represented by a fossil. In paleontology, higher taxonomic surrogacy can provide a more accurate estimate of species richness than can the number of species recorded as fossils.

The effectiveness of higher taxonomic surrogacy greatly depends on the taxonomic level used to estimate species richness. The number of higher taxa does not increase at the same rate as lower taxa because species are not equally distributed among higher taxa; most higher taxonomic groups contain few species while a few higher

taxonomic groups contain many species (Gaston & Williams 1993). Therefore choosing the taxonomic level that meets the goals of the study, known as taxonomic sufficiency, is vital for assessing the effectiveness of higher taxonomic surrogacy (Terlizzi et al. 2003; Groc et al. 2010). In my basic models I used three higher taxonomic levels (genus, family, and order) to predict species richness.

Another issue for higher taxonomic surrogacy is the spatial grain of the data used in the analysis. The smallest spatial resolution (grain) used in many studies is greater than 50 square kilometers; however, most managed areas are only a few square kilometers in size. Surrogates that are effective with large spatial grains might be less effective at smaller grains (Mandelik et al. 2007). One study reported changes in the degree of correlation between family, genus, and species richness with changes in spatial grain at continental extents (Larsen and Rahbek 2005). I examined how higher taxonomic surrogacy varied between two spatial grains: blocks (~20-25km^2) and quads (~100km^2).

The effectiveness of higher taxonomic surrogacy might depend on the region where the study takes place. A study looking at stream macroinvertebrate communities suggested the effectiveness of higher taxonomic surrogacy can differ greatly between regions (Heino 2013). Cross-taxon surrogacy (i.e., using the richness of one taxonomic group to predict the richness of a different taxonomic group) might also be influenced in part by the region of the study (Hess et al. 2006). Regions can differ substantially in the number of species that are present and subsequently affect the ratio of the number of

species to the number of higher taxa. The number of species and higher taxa could differ between regions because urban development is more prevalent in one region and less prevalent in another region. Changes in climate from one region to another could also affect the number of species and higher taxa. If the ratio of species to higher taxa is low (i.e., the number of species closely matches the number of higher taxa) I expect higher taxonomic surrogacy to function well, if the ratio of species to higher taxa is high (i.e., the number of species is vastly different from the number of higher taxa), I would expect higher taxonomic surrogacy to function poorly (Heino 2013).

No studies of higher taxonomic surrogacy have considered the effect of human influence on the relationship between higher taxonomic level richness and species richness. Most threats to biodiversity are due to human population growth, including increasing habitat destruction and fragmentation (Loucks et al. 2008). Human influence could affect higher taxonomic surrogacy because humans have a large impact on biodiversity. I compared the basic surrogacy models to models that included human influence. Every additional variable added to the higher taxonomic surrogacy model will improve the predictive ability of the model. With this study, I was not as interested in comparing predictive ability between the models; rather I was interested in whether the human influence models are significantly different from the basic surrogacy models. The objectives of my study were to examine how higher taxonomic surrogacy is affected by 1) taxonomic level, 2) grain size, 3) ecoregion, and 4) human influence.

METHODS

Atlas Data

I used breeding bird atlases (BBAs) from six states: Colorado 1987-1994 (Kingery 1998); Florida 1986-1991 (Florida Fish & Wildlife 2003); Michigan 1983-1988 (Brewer et al. 1991); New York 1980-1985 (Andrle et al. 1988); Pennsylvania 1983-1989 (Brauning 1992); and Washington 1987-1996 (Smith et al. 1997). These are surveys of the breeding birds within each state, covering a time period of four to seven years. These state's BBAs were chosen based on availability of data, similarity of methods, and sampling coverage within each state. New York was divided into quads approximately 100km^2 and further divided into blocks (four blocks per quad) approximately 25km^2. The other five states (Colorado, Florida, Michigan, Pennsylvania, and Washington) were divided into quads (based on 7.5 minute topographic quads) of approximately 120km^2 and further divided into blocks (six blocks per quad) approximately 20km^2 (Fig. 1). Breeding bird atlases for the Colorado, Florida, Michigan, and Pennsylvania datasets, specified one or more priority blocks within each quad. Sampling effort was standardized within priority blocks while effort varied within non-priority blocks. In a preliminary analysis, priority and non-priority blocks did not differ significantly within the Colorado, Florida, and Pennsylvania atlases so I used both priority blocks and non-priority blocks. Priority blocks and non-priority blocks differed within the Michigan atlas so I only used priority blocks. There were no priority blocks designated within the New York and Washington atlases.

Volunteers surveyed designated blocks for signs of breeding birds. Birds that were unidentifiable or not species specific (e.g., HAWK=Hawk species) were removed from the atlas data. I used the American Ornithologists' Union (AOU) Checklist of North and Middle American Birds (AOU 2013) as the taxonomic authority for this study. Hybrids and other species not recognized by the AOU were excluded from richness calculations. I calculated species richness, genus richness, family richness, and order richness by counting the number of species, genera, families, and orders, respectively, within each atlas block.

Ecoregion Data

An ecoregion is an area classified by ecosystem type. There are three ecoregion levels for North America available from the U.S. Environmental Protection Agency: Level I, Level II, and Level III (EPA 2013). Level I provides a broad classification of continental ecosystems. Level II is a more detailed classification of ecological regions within Level I. Level III further subdivides the ecological regions defined by Level II. I used the Level II Ecoregions of North America for my analysis (Fig. 2). I chose Level II Ecoregions because two to four similarly sized regions occur in each of my selected states. I assigned blocks and quads to an ecoregion based on the location of the center of the block and quad. If the center of a block or quad was located in water or beyond the extent of the ecoregion dataset, they were assigned to the nearest ecoregion.

Human Influence Data

To estimate the influence of humans on biodiversity I adapted the data and

methods (including human influence scores) from The Human Footprint and the Last of the Wild program (Sanderson et al. 2002). Human influence is a composite of several datasets: population density, land transformation, accessibility, and electrical power infrastructure (Table 1).

From the Socioeconomic Data and Applications Center, I obtained 1990 population density from the Gridded Population of the World, v3 dataset (GPWv3 2005). This dataset is an estimate of people per square kilometer within grid cells at 2.5 arc-minute resolution, approximately 5km at the equator. Population density was assigned human influence scores of 0 to 10 with 0 being the smallest amount of human influence and 10 the highest amount of human influence. Population densities of 0 received a score of 0, population densities of 1 were scored as 1, etc.; any density value equal to or greater than 10 people per square kilometer was assigned a score of 10.

I obtained land cover data for 1992 from the Multi-Resolution Land Characteristics Consortium and National Land Cover Database (Vogelmann et al. 2001). This dataset is an unsupervised classification of Landsat Thematic Mapper data at a spatial resolution of 30 m. Thematic Mapper data was classified into sixteen land cover types and each type of landcover was assigned a value between 0 and 10 based on the estimated human influence. I scored land cover classifications including Open Water, Perennial Snow/Ice, Deciduous Forest, Evergreen Forest, Mixed Forest, Shrub/Scrub, Herbaceous, Wood Wetlands, and Emergent Herbaceous Wetlands as 0. Developed Open Space and Developed Low Intensity were classified as 7. Areas classified as

Developed Medium Intensity, Barren Land, and Hay/Pasture were scored as 8. I scored areas classified as Developed High Intensity and Cultivated crops were scored the highest values of 10.

I obtained data layers for built up centers (i.e., large cities), roads, railways, and coastlines, from the Vector Map Level 0 dataset (NIMA 1997). I assigned built up centers a human influence score of 10. Roads, railways, and coastlines were used to estimate human access to an area. I assigned areas within 2 km of these access points a score of 8. Areas between 2 and 15 km were assigned a score of 4. Areas greater than 15 km from roads, railways, and coastlines were assigned a human influence score of 0.

I obtained the stable lights dataset (Version 4 Nighttime Lights) for 1992 from the National Geophysical Data Center website (NOAA 2013). Stable lights are lights which are observable for several hours nearly every night. This dataset is a composite of stable lights for one year and was recorded at a spatial resolution of 30 arc seconds or ~1 kilometer at the equator. I classified stable lights into four categories. Areas with no recorded lights were assigned a human influence score of 0. Lights observed 1% to 40% of nights were assigned a score of 4. Lights observed on 40% to 88% of nights were scored 8. Lights that were observed on more than 89% of nights were scored as 10, signifying the greatest human influence.

The scored data layers were added together in ArcMap10 (ESRI 2011) to create a map that estimated human influence within each state. The mean human influence was calculated with zonal statistics for blocks and quads for the six BBA (Fig. 3). Blocks

and quads that occurred outside of the area scored for human influence and were removed from further analysis.

Analyses

I used regression models to examine how grain, region, and human influence affect higher taxonomic surrogacy. I used the statistical program R (version 2.15.2) to develop the regression models for this study (R 2012). Outliers were identified and removed within R using the aq.plot procedure. Assumptions for all regressions were assessed graphically. A significance level of 0.05 was used for all statistical comparisons.

I used data from blocks to develop three basic surrogacy models that used higher taxonomic levels (genus richness, family richness, and order richness) to predict species richness (Table 2). I then compared the R^2 values for these three basic surrogacy models to determine which higher taxonomic level best predicted species richness.

For the grain size analyses, I developed surrogacy models for each of the higher taxonomic levels that included blocks (small grain size) and quads (large grain size) (Table 2). The remainder of my analyses used richness data from blocks. I used Student's t test (Zar 1998) to compare the higher taxon coefficients (slope) between models generated with block richness and models generated with quad richness. I placed greater emphasis on higher taxon coefficients rather than R^2 values for comparing models. The R^2 value is a measurement of how well the model fits the data but it does not always indicate a difference between models if such a difference exists. Models can

have the same R^2 value (e.g., R^2=0.872) while their coefficients might be very different (e.g., one model has a negative coefficient and the other model has a positive coefficient). I also compared the R^2 values of the block and quad models.

To examine the affect of region on higher taxonomic surrogacy, I developed surrogacy models for each of the ecoregions in each state (Table 2). I compared the higher taxonomic coefficients of these ecoregion models using t tests. Most of the states had three or more t test comparisons between ecoregions so I used the BY-FDR method to correct for Type I error inflation resulting from multiple tests (Narum 2006). I also compared the higher taxon coefficients from the ecoregion models to higher taxon coefficients from the basic models to test if the ecoregion models were significantly different from the state models.

To test if human influence has an effect on higher taxonomic surrogacy, I developed surrogacy models that used higher taxonomic richness and human influence to predict species richness (Table 2). The higher taxon coefficients of these models were compared to those of the basic surrogacy models by using a t test. A significant difference between these coefficients would indicate that human influence affected the surrogacy models.

RESULTS

Basic Surrogacy Model

Before I could examine the effect of grain size, region, and human influence on higher taxonomic surrogacy, I first needed to establish that higher taxa richness could indeed be used to predict species richness with these datasets. All of the basic models were statistically better than random at predicting species richness (Table 3, $p<0.001$). Genus, family, and order richness were always positively associated with species richness (i.e., as the number of species increased, the number of genera, families, and orders also increased). In every state the higher taxon coefficients (slope) increased in value as the taxonomic level increased (Table 3).

In all states, genus richness was an effective surrogate for species richness (Table 3, R^2=0.884 to 0.986). Family richness was an acceptable surrogate in most states (Table 3, R^2=0.791 to 0.910) though less effective in New York (Table 3, R^2=0.691, df=4915, $p<0.001$). Order richness was not as good of a surrogate for species richness in any of the states (Table 3, R^2=0.565 to 0.673). These results suggest genus and family richness can be used to predict species richness, while order richness be used with caution.

Grain Size

I investigated the effect of grain size (blocks and quads) on the basic surrogacy model by comparing higher taxonomic surrogacy models generated with the small grain size (blocks) and the large grain size (quads). All of the surrogacy models were statistically better than random at predicting species richness (Table 4, $p<0.001$). In

every state and grain size, the higher taxon coefficients (slope) increased in value as the taxonomic level increased (Table 4). In four of six states the higher taxonomic coefficients were larger in quad models than in block models (Table 4). Colorado and Pennsylvania were the only exceptions. In Colorado, quad model coefficients were less than block model coefficients when genus and order richness were used to predict species richness (Table 4). In Pennsylvania, quad model coefficients were less than block model coefficients when order richness was used to predict species richness (Table 4). Typically, lower taxonomic levels were better at predicting species richness than higher taxonomic levels (i.e., family richness was better than order richness at predicting species richness). However, within New York and Pennsylvania, order richness within blocks had a slightly larger R^2 value than family richness within quads (Table 4, order richness within blocks R^2= 0.565, 0.599; family richness within quads R^2= 0.5553, 0.5883).

The higher taxon coefficients of the block and quad models were significantly different from each other for most of the comparisons (Table 4). This indicates grain size had a significant affect on higher taxonomic surrogacy models. The two exceptions to this pattern were Colorado and Florida when order richness was used to predict species richness (Table 4, Colorado t=0.444, df=409, p=0.361; Florida t=1.890, df= 4193, p=0.067). Thus, when using order as a surrogate for species richness within these two states, grain size did not affect surrogacy. Generally, grain size significantly influenced the basic surrogacy models prediction of species richness for higher taxonomic levels (genus richness, family richness, and order richness). This suggests that grain size has a

significant effect on higher taxonomic surrogacy.

Ecoregion

I compared surrogacy models from each ecoregion within the state and compared these ecoregion models to the basic surrogacy model for the entire state. For each state I also compared the ecoregion surrogacy models to each other. All ecoregion models were statistically better than random at predicting species richness (Table 5, $p<0.001$). Genus richness continued to be the best predictor of species richness within the ecoregion models (Table 5, R^2=0.898 to 0.991). However, depending on the ecoregion, the predictive ability of other higher taxon levels (family and order) varied within each state (Table 5). For example, within Colorado ecoregions the predictive ability of order richness varied greatly: Western Cordillera R^2=0.635, South Central Semi-arid Prairies R^2=0.810, and Cold Deserts R^2=0.549).

When I compared the ecoregion surrogacy model coefficients to the basic surrogacy model coefficients for the same state and same higher taxon levels, approximately half of the comparisons differed significantly from each other (Table 6). For each state I also compared the ecoregion surrogacy model coefficients between all of the ecoregions. The ecoregion comparisons indicated that higher taxon coefficients differed significantly between ecoregions within the same state most of the time (Table 7). In Florida, ecoregions significantly affected all of the higher taxonomic surrogacy model comparisons (Table 7). These results suggest that the region from which surrogacy models are generated can have a significant effect on higher taxonomic

surrogacy.

Human Influence

I examined if human influence affected higher taxonomic surrogacy. For each state, I compared higher taxon coefficients from models that included human influence to the basic surrogacy model coefficients (without human influence). All of the models including human influence were significantly better than random at predicting species richness (Table 8, $p<0.001$). Higher taxonomic surrogacy models that included human influence appeared to improve the R^2 value only slightly over the models that did not include human influence (Table 8). When I looked at the human influence coefficients (Table 2) I discovered that in most cases (16 out of18) there was an inverse relationship between human influence and species richness (i.e., areas with higher human influence had lower species richness); however, Pennsylvania had a positive relationship between human influence and species richness (i.e., areas with higher human influence had higher species richness; Table 8). Pennsylvania's genus and family richness models showed that species richness was positively associated with human influence (Table 8, genus model human influence coefficient=0.046, df=4785,$p<0.001$ and family model human influence coefficient=0.119, df=4785, $p<0.001$).

Within the most states, higher taxon coefficients from models including human influence did not differ significantly from higher taxon coefficients from the basic surrogacy models (Table 8). New York was the only state in which the models including human influence differed significantly from the basic surrogacy models (Table 8, t=4.308

to 9.920, df=9599, $p<0.001$). These data indicate that while the addition of human influence increased the R^2 value from the basic surrogacy model, human influence did not significantly affect higher taxonomic surrogacy for most of the states analyzed.

DISCUSSION

Higher taxonomic surrogacy could be a useful tool for conservation if used correctly. Gaston and Williams (1993) suggested higher taxonomic surrogacy could reduce the cost and amount of time needed for biological assessments. Higher taxonomic surrogacy could be valuable in citizen science programs due to the reduced amount of training required to identify organisms to genus or family level. It could also be used to rapidly assess areas that have not been studied to locate biodiversity hotspots. While numerous studies have investigated higher taxonomic surrogacy, few have studied the effect of grain and region on this surrogacy technique. I am aware of none that incorporate human influence. These variables all have the potential to affect higher taxonomic surrogacy.

Taxonomic Level

Higher taxonomic surrogacy relies on the predictive relationship between higher taxa richness and species richness (e.g, genus richness, family richness, and order richness). According to these results, higher taxonomic surrogacy worked best when genus richness was used to predict species richness. Family richness was still a good predictor of species richness while order richness was a poor surrogate for species richness. The higher taxon coefficients (slope) increased as the taxonomic level increased because the number of species remained the same while the number of genera, families, and orders decreased. There are fewer higher taxa than species in the taxonomic

hierarchy; therefore there will be a greater difference and variation between species richness and higher taxa richness at the higher taxonomic levels (i.e, predictive power decreased as the higher taxonomic level increased). Previous studies have reported different results regarding the predictive power of higher taxonomic levels (Balmford 2000; Villaseñor 2004; Mazaris 2008; Mazaris 2010; Kallimanis 2012). Some studies that investigated higher taxonomic surrogacy for birds, mammals, amphibians, and reptiles, reported genus and family are good predictors of species richness (Mazaris 2008; Kallimanis 2012). One study of higher taxonomic surrogacy for plants reported genus, family, and order richness were all effective surrogates (Villaseñor 2004), while another study reported only genus richness was effective (Mazaris 2010). Yet another study evaluated the effectiveness of higher taxonomic surrogacy for macrofungi and reported genus richness was the best predictor of species richness, but family and order richness were not effective (Balmford 2000). It is clear from my results and the results of previous research that the effectiveness of higher taxonomic surrogacy depends on the taxonomic level used as a surrogate for species richness. My results indicate that genus richness and family richness were good surrogates for species richness in breeding birds while order richness was less useful.

Grain

Grain size refers to the spatial resolution of the data and should be selected based on the study requirements. I studied how grain size affected higher taxonomic surrogacy. My results suggest that grain size influences higher taxonomic surrogacy. Higher taxon

coefficients were usually significantly larger in quad models (large grain) than in block models (small grain). In most instances, grain size affected higher taxonomic surrogacy. My findings agree with a similar surrogacy study that investigated the relationship between species richness of one taxonomic group (e.g., amphibians) and the combined species richness of the remaining taxonomic groups (e.g., birds, butterflies, freshwater fish, mammals, freshwater mussels, and reptiles; Hess et al. 2006). They reported that as grain size increased, correlation strength increased between mammal richness and the remaining taxa, and between fish richness and the remaining taxa (Hess et al. 2006). When mussel richness and reptile richness were compared to the remaining taxonomic groups, correlation strength decreased as grain size increased (Hess et al. 2006. In my study, the coefficients of determination (R^2) decreased as grain size increased. Larsen & Rahbek reported that grain size was a significant factor when higher taxonomic surrogacy was used for continental conservation priority setting (2005). In this study, higher taxonomic surrogacy performed best for conservation priority-setting when the largest grain size ($\geq 4°$) was used.

Higher taxonomic richness always better predicted species richness at the small grain size (blocks). I suspect this pattern occurred in the breeding bird atlas datasets because richness calculated for the large grain size (quads) could include more genera, families, or orders than the small grain size. Additional higher taxonomic levels (i.e., genera, families, orders) that were species poor could affect higher taxonomic surrogacy. For instance, the large grain size covers a larger spatial extent and could thus include a

number of species poor orders. With each additional species poor order, the number of new species could increase by one. There might be differences in the way grain size affects higher taxonomic surrogacy across different taxonomic groups (i.e., birds, mammals) but I am not aware of any study that that has investigated this issue.

Studies of higher taxonomic surrogacy have been completed at spatial grains larger than or the same size as my study, however management decisions often need to be made at much smaller spatial grains. Future research should compare several taxonomic groups at grain sizes that will benefit local conservation decisions. My results and the results of previous studies suggest grain size is a vital factor to consider when evaluating the effectiveness of higher taxonomic surrogacy; models developed at one spatial grain should not be applied to different spatial grains.

Ecoregion

The number of species, genera, families, and orders recorded in an area is dependent on the geographic location of that area, therefore I studied the affect of ecoregion on higher taxonomic surrogacy. The results of my analyses suggest ecoregion has a significant effect on higher taxonomic surrogacy. Models based on ecoregion differed from basic surrogacy models that included the entire state about half of the time. A previous study compared cross-taxon surrogacy (i.e., taxa richness from one group is used to predict species richness of one or more other taxonomic groups) between ecoregions for several taxa (e.g. amphibians, birds, butterflies, freshwater fish, mammals, freshwater mussels, and reptiles) in the Mid-Atlantic and Pacific Northwest (Hess et al.

2006). This study recommended that cross-taxa surrogacy models should not be used in a region other than the one for which it was developed. This study agreed with my results suggesting that region can affect surrogacy. This can be explained by the variation in biodiversity between ecoregions. When states were divided into ecoregions, each ecoregion contained a subset of the data from the entire state. Therefore, each ecoregion can vary in the richness of species, genera, families, and orders.

I think higher taxonomic surrogacy models generated by state potentially can be used to predict species richness for the entire state or within smaller regions of that state. Models developed for the entire state will still encompass the smaller regions, as long as data from that region were used in developing the model for the entire state. However, a higher taxonomic surrogacy model developed within a single ecoregion should not be applied to other regions of the state or to the state as a whole because that ecoregion model would not include data from other regions of the state.

Human Influence

There is no location on Earth that has not been directly or indirectly influenced by humans. This was why I investigated the affect of human influence on higher taxonomic surrogacy. My results suggest human influence does not greatly affect higher taxonomic surrogacy. It is possible that the geographical extent of my study did not capture the effect of human influence. This could be tested by comparing models with and without human influence at several spatial extents (i.e., extents both smaller and larger than the ecoregion and state extents used within this study). It is also possible that some of the

data layers I used to calculate the total human influence might affect biodiversity more (or differently) than other layers. For example, built up areas might influence biodiversity more than human access to an area by rivers. It might be beneficial to separate the data layers from the human influence dataset and develop models with each layer individually.

Usually human influence exhibited an inverse relationship to species richness (i.e., as human influence increased, species richness decreased). There were two instances where human footprint had a positive association with species richness (i.e., as human influence increased, species richness increased). Ecologically this could be the result of natural species richness gradients (i.e., areas with naturally high species richness despite higher human influence while depauperate areas have lower human influence). However, the positive relationship between species richness and human influence in these two instances is likely not biologically meaningful because the human influence coefficients are very small and could be the result of a Type I error.

Additionally, although the inclusion of human influence to the higher taxonomic surrogacy model slightly increased the R^2 value, the addition of any independent variable to a model will always increase the model's R^2 value. My results did not focus on the R^2 value, however other tests (e.g., Akaike's Information Criterion) might indicate the best model (i.e., model with human influence versus model without human influence) based on goodness of fit and parsimony. While human influence can affect biodiversity, my results suggest that human influence does not appear to significantly affect higher

taxonomic surrogacy when applied to breeding bird atlas data.

Additional Research Recommendations

Another factor not examined here that might influence the effectiveness of higher taxonomic surrogacy is the taxonomic group (e.g., birds) used within this study. Studies of higher taxonomic surrogacy have been conducted for a variety of taxa including fungi, birds, and plants (Balmford 2000; Mazaris 2010; Kallimanis 2012). Strong relationships between higher taxa and species richness have been demonstrated for various taxonomic groups especially between genus richness and species richness. However, the relationship for other higher taxonomic levels (family and order) varies for different taxonomic groups (Balmford 2000; Villaseñor 2004; Mazaris 2008; Mazaris 2010; Kallimanis 2012). While my results demonstrate that other factors (e.g., grain, region) also can influence the surrogacy relationship, additional research should study how the taxonomic group affects higher taxonomic surrogacy.

Comparison of the effectiveness of higher taxa surrogacy between areas with native taxa and areas with non-native or invasive taxa could be informative. Invasive species are the second greatest threat to biodiversity (Wilcove 1998). I expect that surrogacy might be more effective in areas with larger numbers of invasive organisms, especially at smaller spatial scales. At smaller spatial scales invasive plant species can create stands of monocultures. If only one or two species exist in an area, higher taxonomic richness could be a very accurate predictor of species richness. This relationship might be negligible at larger scales.

Many factors might influence higher taxonomic surrogacy. I have highlighted some aspects that require additional research; however, there are other factors that I did address in this study (e.g., how evolutionary relationships are reflected in taxonomic rankings). Gaston and Williams (1993) highlighted several factors that had known or potential affects on higher taxonomic surrogacy. There are still aspects of higher taxonomic surrogacy that should be investigated to ensure this technique is used effectively.

Implications for Conservation

If higher taxonomic surrogacy is to be used for conservation planning, we must consider the taxonomic level, grain, and region (ecoregion) used to develop the surrogacy model. Taxonomic levels above genus should be used with caution, as family and order were less reliable when predicting species richness. Also, the spatial grain for which the higher taxonomic surrogacy model is used for conservation should be similar to the spatial grain used when the model was developed. Region can affect higher taxonomic surrogacy; therefore, caution should be exercised when surrogacy relationships are applied to regions other than those for which it was developed. According to my results, human influence does not usually affect higher taxonomic surrogacy. Further research should be done to look at the effect of human influence on higher taxonomic surrogacy as it might affect other taxonomic groups differently. Additionally, although I did not make statistical comparisons among states, the predictive ability of family and order richness varied between states (i.e., the R^2 values varied depending on the state). This suggests

factors such as sampling effort or species pool also might influence higher taxonomic surrogacy. Additional research should examine the affects of sample effort on higher taxonomic surrogacy.

I suggest that higher taxonomic surrogacy be used cautiously, especially when taxonomic levels higher than genus are used for prediction. I think the greatest potential value of higher taxonomic surrogacy is to quickly assess biodiversity in an area for conservation prioritization, thereby saving time and money. Conservation biologists face the daunting task of to preserving biodiversity within a landscape that will continue to change rapidly as the human world population approaches eight billion (United States Census Bureau 2014). Finding efficient methods to monitor biodiversity is a high priority and I believe higher taxonomic surrogacy might be an effective way to estimate biodiversity.

LITERATURE CITED

Allmon, W. D. 1992. Genera in Paleontology: Definition and Significance. Historical Biology **6**:149-158.

Andelman, S. J., and F. F. Fagan. 2000. Umbrellas and flagships: Efficient conservation surrogates or expensive mistakes? Proceedings of the National Academy of Sciences **11**:5954-5959.

Andrle, R.F., and J. F. Carroll. 1988. The Atlas of Breeding Birds in New York State. Cornell University Press, Ithaca, New York.

AOU Checklist. 2013. AOU Checklist of North and Middle American Birds. American Ornithologists' Union. Available from http://checklist.aou.org/taxa/ (accessed April 2013).

Balmford, A., A. J. E. Lyon, and R. M. Lang. 2000. Testing the higher-taxon approach to conservation planning in a megadiverse group: the macrofungi. Biological Conservation **93**:209-217.

Barnosky, A. D., N. Matzke, S. Tomiya, G. O. U. Wogan, B. Swartz, T. B. Quental, C. Marshall, J. L. McGuire, E. L. Lindsey, K. C. Maguire, B. Mersey and E. A. Ferrer. 2011. Has the Earth's sixth mass extinction already arrived? Nature **471**:51-57.

Brauning, D. W. 1992. Atlas of breeding birds in Pennsylvania. University of Pittsburgh Press, Pittsburgh, Pennsylvania.

Brewer, R., G. A. McPeek, and R. J. Adams Jr. 1991. The atlas of breeding birds of Michigan. Michigan State University Press, East Lansing, Michigan.

Caro, T. M. and G. O'Doherty. 1999. On the use of surrogate species in conservation biology. Conservation Biology **13**:805–814.

Caro, T. M. 2003. Umbrella species: critique and lessons from East Africa. Animal Conservation **6**:171–181.

EPA. 2013. Ecoregion Level II. U. S. Environmental Protection Agency: Western Ecology Division Available from http://www.epa.gov/wed/pages/ecoregions/na_eco.htm#Level II (accessed February 2014).

ESRI. 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

Florida Fish & Wildlife Conservation Commission. 2003. Florida's breeding bird atlas: A collaborative study of Florida's birdlife. Available from http://legacy.myfwc.com/bba/data/default.asp (accessed February 2013).

Gaston, K. J., and P. H. Williams. 1993. Mapping the World's Species-The Higher Taxon Approach. Biodiversity Letters **1**:2-8.

GPWv3: Gridded Population of the World, Version 3 Population Density Grid. 2005. Center for International Earth Science Information Network (CIESIN) Columbia University.New York, New York. Available from

http://sedac.ciesin.columbia.edu/data/set/gpw-v3-population-density (accessed April 2013).

Groc, S., J. H. C. Delabie, J. T. Longino, J. Orivel, J. D. Majer, H. L. Vasconcelos, and A. Dejean. 2010. A new method based on taxonomic sufficiency to simplify studies on Neotropical ant assemblages. Biological Conservation **143**:2832-2839.

Heino, J. 2013. Taxonomic surrogacy, numerical resolution and responses of stream macroinvertebrate communities to ecological gradients: Are the inferences transferable among regions? Ecological Indicators **35**:186-194.

Hess, R. G., R. A. Bartel, A. K. Leidner, K. M. Rosenfeld, M. J. Rubino, S. B. Snider, and T. H. Ricketts. 2006. Effectiveness of biodiversity indicators varies with extent, grain, and region. Biological Conservation **132**:448-457.

Jones, F. C. 2008. Taxonomic sufficiency: the influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. Environmental Reviews **16**:45-69.

Kallimanis, A. S., A. D. Mazaris, D. Tsakanikas, P. Dimopoulos, J. D. Pantis, and S. P. Sgardelis. 2012. Efficient biodiversity monitoring: Which taxonomic level to study? Ecological Indicators **15**:100-104.

Kingery, H. E. 1998. Colorado breeding bird atlas. Colorado Bird Atlas Partnership.

Landres P. B., J. Verner, and J. W. Thomas. 1988. Ecological Uses of Vertebrate Indicator Species: A Critique. Society for Conservation Biology **2**:316-328.

Larsen, F. W., and Rahbek, C. 2005. The influence of spatial grain size on the suitability of the higher-taxon approach in continental priority-setting. Animal Conservation **8**:389-396.

Loucks, C., T. H. Ricketts, R. Naidoo, J. Lamoreux, and J. Hoekstra. 2008. Explaining the global pattern of protected area coverage: relative importance of vertebrate biodiversity, human activities and agricultural suitablility. Journal of Biogeography **35**:1337-1348.

Mandelik, Y., T. Dayan, V. Chikatunov, and V. Kravchenko. 2007. Reliability of a Higher-Taxon Approach to Richness, Rarity, and Composition Assessments at the Local Scale. Conservation Biology **21**:15-1515.

Marshall, J. C., A. L. Steward, and B. D. Harch. 2006. Taxonomic Resolution and Quantification of Freshwater Macroinvertebrate Samples from an Australian Dryland River: The Benefits and Costs of Using Species Abundance Data. Hydrobiologia **572**:171-194.

Mazaris A. D., A. S. Kallimanis, S. P. Sgardelis, J. D. Pantis. 2008. Does higher taxon diversity reflect richness of conservation interest species? The case for birds, mammals, amphibians, and reptiles in Greek protected areas. Ecological Indicators **8**:664-671.

Mazaris, A. D., A. S. Kallimanis, J. Tzanopoulos, S. P. Sgardelis, and J. D. Prantis. 2010. Can we predict the number of plant species from the richness of a few common genera, families, or orders? Journal of Applied Ecology **47**:662-670.

Mora, C., D. P. Tittensor, S. Adl, A. G. B. Simpson, and B. Worm. 2011. How many species are there on Earth and in the ocean? PLos Biology **9**:1-8.

Narum, S. R. 2006. Beyond Bonferroni: Less conservative analyses for conservation genetics. Conservation Genetics **7**:783-787.

NIMA: National Imagery and Mapping Agency. 1997. Vector Map Level 0 (VMAP0), ed. 003. Washington DC: NIMA. Available from http://geoengine.nima.mil/ftpdir/archive/vpf_data/v0noa.tar.gz (accessed August 2013)

NOAA: National Oceanic and Atmospheric Administration. 2013. Version 4 DMSP-OLS Nighttime Lights Time Series: Average Visible, Stable Lights, & Cloud Free Coverages F10 year 1992. Available from http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html (accessed August 2013).

Noss, R. F. 1990. Indicators for monitoring biodiversity: a hierarchical approach. Conservation Biology **4**:355–364.

Pimm, S. L., G. J. Russel, J. L. Gittleman, and T. M. Brooks. 1995. The future of biodiversity. Science **269**:347-350.

R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org/ (accessed Octorber 2012).

Roberge, J. M., and P. Angelstam. 2004. Usefulness of the umbrella species concept as a conservation tool. Conservation Biology **18**:76-85.

Rubinoff, D. 2001. Evaluating the California gnatcatcher as an umbrella species for conservation of southern California coastal sage scrub. Conservation Biology **15**:1374-1383.

Sanderson, E. W., M. Jaiteh, M. A. Levy, K. H. Redford, A. V. Wannebo, and G. Woolmer. 2002. The Human Footprint and the Last of the Wild. BioScience **52**:891-904.

Signor, P. W. 1990. The geologic history of diversity. Annual Review of Ecolology and Systematics **21**:509-539.

Smith, M. R., P. W. Mattocks, and K. M. Cassidy. 1997. Breeding birds of Washington State: Location data and predicted distributions (No. 1). Seattle Audubon Society.

Stanislao, B., A. Terlizzi, J. Claudet, S. Fraschetti, and F. Boero. 2012. Taxonomic relatedness does not matter for species surrogacy in the assessment of community responses to environmental drivers. Journal of Applied Ecology **49**:357-366.

Stork, N. E. 1993. How many species are there? Biodiversity and Conservation **2**:215-232.

Terlizzi, A., S. Bevilacqua, S. Fraschetti, F. Boero. 2003. Taxonomic sufficiency and the increasing insufficiency of taxonomic expertise. Marine Pollution Bulletin **46**:556-561.

Tracy, C. R., and P. F. Brussard. 1994. Preserving biodiversity: species in landscapes. Ecological Applications **4**:205-206.

United States Census Bureau. 2014. U.S. and World Population Clock, Suitland, Maryland. Available from https://www.census.gov/popclock/ (accessed March 2014).

Villaseñor, J. L., Ibarra-Manríquez G., Meave J. A., and Ortíz E. 2004. Higher Taxa as Surrogates of Plant Biodiversity in a Megadiverse Country. Conservation Biology **19**:232-238.

Vogelmann, J.E., S.M. Howard, L. Yang, C. R. Larson, B. K. Wylie, and J. N. Van Driel. 2001. Completion of the 1990's National Land Cover Data Set for the conterminous United States, Photogrammetric Engineering and Remote Sensing **67**:650-662.

Warwick, R. M., and K. R. Clarke. 1995. New 'biodiversity' measures reveal a decrease in taxonomic distinctness with increasing stress. Marine Ecology Progress Series **129**:301-05.

Western, D. 1987. Africa's elephants and rhinos: flagships in crisis. Trends in Ecology **2**:343-346.

Wilcove D. S., D. Rothstein, J. Dubow, A. Phillips, and E. Losos. 1998. Quantifying Threats to Imperiled Species in the United States. BioScience **48**:607-615.

Williams, P. H., N. D. Burgess, and C. Rahbek. 2000. Flagship species, ecological complementarity and conserving the diversity of mammals and birds in sub-Saharan Africa. Animal Conservation **3**:249-260.

Wilson, E. O. 1992. The Diversity of Life. Page 243. Belknap Press, Cambridge, Massachusetts.

Zar, J. H. 1998. Biostatistical Analysis. Page 718 in 4th Edition. Prentice Hall, Upper Saddle River, New Jersey.

Table 1. List of data used to produce the human influence dataset.

Dataset type	*Dataset name*	*Year*	*Sources*
Population Density	Gridded Population of the World (GPW), v3: Population Density Grid	1990	CIESIN[a]/SEDAC[b]
Land transformation	Land Cover	1992	MRLC[c]/NLCD[d]
	Vector Map Level 0 Built-Up Centers	1960's-1990's	NIMA[e]
	Vector Map Level 0 Roads		
	Vector Map Level 0 Railways		
Access	Vector Map Level 0 Roads	1960's-1990's	NIMA[e]
	Vector Map Level 0 Railways		
	Vector Map Level 0 Rivers		
	Vector Map Level 0 Coastlines		
Electrical power	Version 4 Nighttime Lights: Average Visible, Stable Lights, & Cloud Free Coverage	1992	NOAA[f]/NGDC[g]

[a]CIESIN, Center for International Earth Science Information Network; [b]SEDAC, Socioeconomic Data and Applications Center; [c]MRLC, Multi-Resolution Land Characteristics Consortium; [d]NLCD, National Land Cover Database; [e]NIMA, National Imagery and Mapping Agency; [f]NOAA, National Oceanic and Atmospheric Administration; [g]NGDC, National Geophysical Data Center

Table 2. Illustration of surrogacy models and coefficients used within analyses.

Model type	*Model*
Basic Surrogacy Model	Species Richness = B_1^a Genus Richness + C
	Species Richness = B_2^a Family Richness + C
	Species Richness = B_3^a Order Richness + C
Grain Size	Block Species Richness = B_1^a Block Genus Richness + C
	Block Species Richness = B_2^a Block Family Richness + C
	Block Species Richness = B_3^a Block Order Richness + C
	Quad Species Richness = B_1^a Quad Genus Richness + C
	Quad Species Richness = B_2^a Quad Family Richness + C
	Quad Species Richness = B_3^a Quad Order Richness + C
Ecoregion	Ecoregion 1 Species Richness = B_1^a Genus Richness+ C
	Ecoregion 1 Species Richness = B_2^a Family Richness+ C
	Ecoregion 1 Species Richness = B_3^a Order Richness+ C
	Ecoregion 2 Species Richness = B_1^a Genus Richness+ C
	Ecoregion 2 Species Richness = B_2^a Family Richness+ C
	Ecoregion 2 Species Richness = B_3^a Order Richness+ C
Human Influence	Species Richness = B_1^a Genus Richness + B_2^b Human Influence + C
	Species Richness = B_2^a Family Richness + B_3^b Human Influence + C
	Species Richness = B_4^a Order Richness + B_5^b Human Influence + C

[a]Higher Taxon Coefficient
[b]Human Influence Coefficient

Table 3. Results from regression analysis for all states showing a comparison between basic surrogacy models.

Breeding Bird Atlas	*Taxon Level*	*Higher Taxon Coefficient*	*Standard error*	R^2	*df*	*p-value*
Colorado	genus	1.182	0.009	0.984	288	<0.001
	family	2.528	0.052	0.891	288	<0.001
	order	4.332	0.195	0.630	288	<0.001
Florida	genus	1.136	0.002	0.986	3335	<0.001
	family	1.935	0.012	0.891	3335	<0.001
	order	3.475	0.045	0.642	3335	<0.001
Michigan	genus	1.209	0.003	0.979	2805	<0.001
	family	2.907	0.025	0.823	2805	<0.001
	order	4.810	0.063	0.673	2805	<0.001
New York	genus	1.222	0.006	0.884	4915	<0.001
	family	2.837	0.027	0.691	4915	<0.001
	order	4.235	0.053	0.565	4915	<0.001
Pennsylvania	genus	1.269	0.003	0.974	4815	<0.001
	family	3.227	0.024	0.791	4815	<0.001
	order	6.207	0.073	0.599	4815	<0.001
Washington	genus	1.186	0.005	0.982	1243	<0.001
	family	2.280	0.020	0.910	1243	<0.001
	order	4.212	0.092	0.625	1243	<0.001

Table 4. Results from t tests comparing the higher taxon coefficients between grain sizes (blocks and quads) within each state.

			Results from regression					*Results from t test*			
Breeding Bird Atlas	*Taxon Level*	*Grain*	*Higher Taxon Coefficient*	*Standard Error*	R^2	*df*	*p-value*	*t*	*df*	*p-value*	
Colorado	genus	BLOCK	1.182	0.009	0.984	288	<0.001	2.729	409	<0.001	*
		QUAD	1.249	0.023	0.960	123	<0.001				
	family	BLOCK	2.528	0.052	0.891	288	<0.001	3.272	409	0.002	*
		QUAD	3.046	0.150	0.770	123	<0.001				
	order	BLOCK	4.332	0.195	0.630	288	<0.001	0.444	409	0.361	
		QUAD	4.137	0.394	0.468	123	<0.001				
Florida	genus	BLOCK	1.136	0.002	0.986	3335	<0.001	9.695	4193	<0.001	*
		QUAD	1.212	0.007	0.968	860	<0.001				
	family	BLOCK	1.935	0.012	0.891	3335	<0.001	3.616	4193	<0.001	*
		QUAD	2.105	0.046	0.713	860	<0.001				
	order	BLOCK	3.475	0.045	0.642	3335	<0.001	1.890	4193	0.067	
		QUAD	3.781	0.156	0.407	860	<0.001				
Michigan	genus	BLOCK	1.209	0.003	0.979	2805	<0.001	12.389	3770	<0.001	*
		QUAD	1.326	0.009	0.959	967	<0.001				
	family	BLOCK	2.907	0.025	0.823	2805	<0.001	6.429	3770	<0.001	*
		QUAD	3.337	0.062	0.750	967	<0.001				
	order	BLOCK	4.810	0.063	0.673	2805	<0.001	4.559	3770	<0.001	*
		QUAD	5.461	0.128	0.653	967	<0.001				

*denotes significant difference between grain sizes

Table 4. continued

			Results from regression					*Results from t test*			
Breeding Bird Atlas	*Taxon Level*	*Grain*	*Higher Taxon Coefficient*	*Standard Error*	R^2	*df*	*p-value*	*t*	*df*	*p-value*	
New York	genus	BLOCK	1.222	0.006	0.884	4915	<0.001	3.571	6178	<0.001	*
		QUAD	1.291	0.018	0.798	1265	<0.001				
	family	BLOCK	2.837	0.027	0.691	4915	<0.001	8.111	6178	<0.001	*
		QUAD	3.604	0.091	0.555	1265	<0.001				
	order	BLOCK	4.235	0.053	0.565	4915	<0.001	4.641	6178	<0.001	*
		QUAD	5.184	0.198	0.352	1265	<0.001				
Pennsylvania	genus	BLOCK	1.269	0.003	0.974	4815	<0.001	9.168	5643	<0.001	*
		QUAD	1.378	0.012	0.945	830	<0.001				
	family	BLOCK	3.227	0.024	0.791	4815	<0.001	4.335	5643	<0.001	*
		QUAD	3.704	0.107	0.588	830	<0.001				
	order	BLOCK	6.207	0.073	0.599	4815	<0.001	3.816	5643	<0.001	*
		QUAD	4.884	0.339	0.199	830	<0.001				
Washington	genus	BLOCK	1.186	0.005	0.982	1243	<0.001	5.925	1726	<0.001	*
		QUAD	1.251	0.010	0.970	485	<0.001				
	family	BLOCK	2.280	0.020	0.910	1243	<0.001	7.957	1726	<0.001	*
		QUAD	2.674	0.045	0.878	485	<0.001				
	order	BLOCK	4.212	0.092	0.625	1243	<0.001	3.679	1726	<0.001	*
		QUAD	4.931	0.172	0.628	485	<0.001				

*denotes significant difference between grain sizes

Table 5. Results of surrogacy models within ecoregions in all states.

Breeding Bird Atlas	*Ecoregion*	*Taxon Level*	*Higher Taxon Coefficient*	*Standard Error*	R^2	*df*	*p-value*
Colorado	6.2 Western Cordillera	genus	1.228	0.016	0.982	102	<0.001
		family	2.642	0.084	0.905	102	<0.001
		order	4.865	0.363	0.635	102	<0.001
	9.4 South Central Semi-arid Prairies	genus	1.171	0.012	0.991	82	<0.001
		family	2.509	0.077	0.927	82	<0.001
		order	4.885	0.259	0.810	82	<0.001
	10.1 Cold Deserts	genus	1.165	0.015	0.984	100	<0.001
		family	2.603	0.104	0.860	100	<0.001
		order	4.622	0.415	0.549	100	<0.001
Florida	8.3 Southeastern USA Plains	genus	1.153	0.007	0.987	380	<0.001
		family	2.200	0.034	0.917	380	<0.001
		order	3.951	0.128	0.715	380	<0.001
	8.5 Mississippi Alluvial and Southeast USA Coastal Plains	genus	1.135	0.003	0.984	2505	<0.001
		family	1.910	0.013	0.891	2505	<0.001
		order	3.470	0.049	0.667	2505	<0.001
	15.4 Everglades	genus	1.078	0.006	0.987	446	<0.001
		family	1.691	0.027	0.900	446	<0.001
		order	3.007	0.084	0.740	446	<0.001

Table 5. continued

Breeding Bird Atlas	*Ecoregion*	*Taxon Level*	*Higher Taxon Coefficient*	*Standard Error*	R^2	*df*	*p-value*
Michigan	5.2 Mixed Wood Shield	genus	1.188	0.010	0.959	610	<0.001
		family	2.710	0.064	0.746	610	<0.001
		order	3.689	0.144	0.519	610	<0.001
	8.1 Mixed Wood Plains	genus	1.227	0.005	0.979	1490	<0.001
		family	2.975	0.035	0.828	1490	<0.001
		order	4.682	0.084	0.676	1490	<0.001
	8.2 Central USA Plains	genus	1.174	0.005	0.988	701	<0.001
		family	2.821	0.041	0.872	701	<0.001
		order	5.231	0.116	0.743	701	<0.001
New York	5.3 Atlantic Highlands	genus	1.215	0.011	0.898	1405	<0.001
		family	2.889	0.041	0.776	1405	<0.001
		order	4.376	0.090	0.626	1405	<0.001
	8.1 Mixed Wood Plains	genus	1.333	0.007	0.918	3323	<0.001
		family	3.020	0.038	0.660	3323	<0.001
		order	4.320	0.070	0.531	3323	<0.001

Table 5. continued

Breeding Bird Atlas	*Ecoregion*	*Taxon Level*	*Higher Taxon Coefficient*	*Standard Error*	R^2	*df*	*p-value*
Pennsylvania	5.3 Atlantic Highlands	genus	1.284	0.006	0.975	990	<0.001
		family	3.255	0.054	0.789	990	<0.001
		order	6.560	0.166	0.613	990	<0.001
	8.1 Mixed Wood Plains	genus	1.259	0.006	0.980	785	<0.001
		family	3.360	0.055	0.825	785	<0.001
		order	7.081	0.208	0.597	785	<0.001
	8.3 Southeastern USA Plains	genus	1.227	0.008	0.979	496	<0.001
		family	3.171	0.071	0.800	496	<0.001
		order	6.902	0.236	0.633	496	<0.001
	8.4 Ozark, Ouachita-Appalachian Forests	genus	1.276	0.004	0.970	2524	<0.001
		family	3.229	0.034	0.786	2524	<0.001
		order	5.965	0.097	0.599	2524	<0.001
Washington	6.2 Western Cordillera	genus	1.212	0.007	0.979	654	<0.001
		family	2.327	0.033	0.885	654	<0.001
		order	4.275	0.129	0.625	654	<0.001
	7.1 Marine West Coast Forest	genus	1.193	0.009	0.985	276	<0.001
		family	2.337	0.042	0.917	276	<0.001
		order	4.089	0.171	0.673	276	<0.001
	10.1 Cold Deserts	genus	1.149	0.007	0.987	309	<0.001
		family	2.174	0.042	0.896	309	<0.001
		order	4.052	0.168	0.652	309	<0.001

Table 6. Results from t tests comparing the coefficients from ecoregions models to the basic surrogacy models.

Breeding Bird Atlas	*Ecoregion*	*Taxon level*	*t*	*df*	*p-value*	
Colorado	10.1	genus	38.191	386	<0.001	*
		family	0.643	386	0.324	
		order	0.633	386	0.326	
	6.2	genus	31.798	388	<0.001	*
		family	1.153	388	0.205	
		order	1.294	388	0.173	
	9.4	genus	43.680	368	<0.001	*
		family	0.204	368	0.391	
		order	1.703	368	0.094	
Florida	15.4	genus	9.148	3779	<0.001	*
		family	8.371	3779	<0.001	*
		order	4.895	3779	<0.001	*
	8.3	genus	2.385	3713	0.023	*
		family	7.388	3713	<0.001	*
		order	3.516	3713	<0.001	*
	8.5	genus	0.268	5838	0.385	
		family	1.409	5838	0.148	
		order	0.075	5838	0.398	
Michigan	5.2	genus	2.004	3413	0.054	*
		family	2.859	3413	0.007	*
		order	7.143	3413	<0.001	*
	8.1	genus	3.124	4293	0.003	*
		family	1.569	4293	0.117	
		order	1.217	4293	0.190	
	8.2	genus	5.826	3504	<0.001	*
		family	1.786	3504	0.081	
		order	3.186	3504	0.003	
New York	5.3	genus	0.554	6318	0.342	
		family	1.051	6318	0.230	
		order	1.349	6318	0.161	
	8.1	genus	11.839	8236	<0.001	*
		family	3.952	8236	<0.001	*
		order	0.965	8236	0.250	

*denotes significant difference between ecoregion models and basic models

Table 6. continued

Breeding Bird Atlas	*Ecoregion*	*Taxon level*	*t*	*df*	*p-value*	
Pennsylvania	5.3	genus	2.100	5803	0.044	*
		family	0.477	5309	0.356	
		order	1.949	5803	0.060	
	8.1	genus	1.398	5598	0.150	
		family	2.208	5598	0.035	
		order	3.972	5598	<0.001	*
	8.3	genus	4.830	5309	<0.001	*
		family	0.747	5309	0.302	
		order	2.816	5309	0.008	
	8.4	genus	1.302	7337	0.171	
		family	0.049	7337	0.398	
		order	1.991	7337	0.055	
Washington	10.1	genus	4.276	1550	<0.001	*
		family	2.269	1550	0.031	*
		order	0.833	1550	0.282	
	6.2	genus	3.113	1895	0.003	*
		family	1.217	1895	0.190	
		order	0.396	1895	0.369	
	7.1	genus	0.713	1517	0.309	
		family	1.216	1517	0.190	
		order	0.632	1517	0.327	

*denotes significant difference between ecoregion models and basic models

Table 7. Results from t tests comparing the coefficients between ecoregions models within each state.

Breeding Bird Atlas	*Grain*	*Ecoregion*	*t-value*	*df*	*p-value*	
Colorado	genus	10.1 vs. 6.2	2.875	200	0.007	*
		10.1 vs. 9.4	0.340	180	0.376	
		6.2 vs. 9.4	2.804	182	0.008	*
	family	10.1 vs. 6.2	0.291	200	0.382	
		10.1 vs. 9.4	0.724	180	0.306	
		6.2 vs. 9.4	1.164	182	0.202	
	order	10.1 vs. 6.2	0.441	200	0.361	
		10.1 vs. 9.4	0.537	180	0.345	
		6.2 vs. 9.4	0.045	182	0.398	
Florida	genus	15.4 vs. 8.3	8.406	824	<0.001	*
		15.4 vs. 8.5	8.722	2949	<0.001	*
		8.3 vs. 8.5	2.465	2883	0.019	*
	family	15.4 vs. 8.3	11.797	824	<0.001	*
		15.4 vs. 8.5	7.342	2949	<0.001	*
		8.3 vs. 8.5	7.962	2883	<0.001	*
	order	15.4 vs. 8.3	6.167	824	<0.001	*
		15.4 vs. 8.5	4.745	2949	<0.001	*
		8.3 vs. 8.5	3.516	2883	0.001	*
Michigan	genus	5.2 vs. 8.1	3.550	2098	0.001	*
		5.2 vs. 8.2	1.259	1309	0.180	
		8.1 vs. 8.2	7.728	2189	<0.001	*
	family	5.2 vs. 8.1	3.629	2098	0.001	*
		5.2 vs. 8.2	1.461	1309	0.137	
		8.1 vs. 8.2	2.858	2189	0.007	*
	order	5.2 vs. 8.1	5.969	2098	<0.001	*
		5.2 vs. 8.2	8.353	1309	<0.001	*
		8.1 vs. 8.2	3.833	2189	<0.001	*
New York	genus	5.3 vs. 8.1	9.120	4726	<0.001	*
	family	5.3 vs. 8.1	2.342	4726	0.026	*
	order	5.3 vs. 8.1	0.490	4726	0.354	

*denotes significant difference between ecoregions

Table 7. continued

Breeding Bird Atlas	*Grain*	*Ecoregion*	*t-value*	*df*	*p-value*	
Pennsylvania	genus	5.3 vs. 8.1	2.727	1773	0.010	*
		5.3 vs. 8.3	5.474	1484	<0.001	*
		5.3 vs. 8.4	1.018	3512	0.238	
		8.1 vs. 8.3	3.070	1279	0.004	*
		8.1 vs. 8.4	2.160	3307	0.039	
		8.3 vs. 8.4	5.273	3018	<0.001	*
	family	5.3 vs. 8.1	1.364	1773	0.157	
		5.3 vs. 8.3	0.944	1484	0.255	
		5.3 vs. 8.4	0.412	3512	0.367	
		8.1 vs. 8.3	2.099	1279	0.044	
		8.1 vs. 8.4	2.026	3307	0.051	
		8.3 vs. 8.4	0.738	3018	0.304	
	order	5.3 vs. 8.1	1.962	1773	0.058	
		5.3 vs. 8.3	1.187	1484	0.197	
		5.3 vs. 8.4	3.098	3512	0.003	*
		8.1 vs. 8.3	0.570	1279	0.339	
		8.1 vs. 8.4	4.871	3307	<0.001	*
		8.3 vs. 8.4	3.676	3018	<0.001	*
Washington	genus	10.1 vs. 6.2	6.181	961	<0.001	*
		10.1 vs. 7.1	3.852	583	<0.001	*
		6.2 vs. 7.1	1.697	928	0.095	
	family	10.1 vs. 6.2	2.866	961	0.007	*
		10.1 vs. 7.1	2.734	583	0.010	*
		6.2 vs. 7.1	0.187	928	0.392	
	order	10.1 vs. 6.2	1.050	961	0.230	
		10.1 vs. 7.1	0.153	583	0.394	
		6.2 vs. 7.1	0.866	928	0.274	

*denotes significant difference between ecoregions

Table 8. Results from t tests comparing the higher taxon coefficients from the basic surrogacy models to the human influence models.

		Models without Human Influence)					*Models with Human Influence*						
Breeding Bird Atlas	*Taxon Level*	*Higher Taxon Coef.*[a]	*SE*	R^2	*df*	*p-value*	*Higher Taxon Coef.*[a]	*Human Influence Coef.*[b]	*SE*	R^2	*df*	*p-value*	
Colorado	genus	1.243	0.011	0.978	305	<0.001	1.242	-0.031	0.011	0.978	304	<0.001	
	family	2.777	0.062	0.866	305	<0.001	2.781	-0.162	0.061	0.870	304	<0.001	
	order	4.790	0.233	0.580	305	<0.001	4.880	-0.320	0.230	0.596	304	<0.001	
Florida	genus	1.142	0.002	0.985	3403	<0.001	1.146	-0.026	0.002	0.985	3402	<0.001	
	family	1.959	0.012	0.884	3403	<0.001	1.972	-0.038	0.012	0.886	3402	<0.001	
	order	3.523	0.046	0.637	3403	<0.001	3.577	-0.059	0.047	0.640	3402	<0.001	
Michigan	genus	1.209	0.003	0.980	2842	<0.001	1.207	-0.088	0.003	0.981	2841	<0.001	
	family	2.904	0.024	0.834	2842	<0.001	2.938	-0.304	0.023	0.853	2841	<0.001	
	order	4.932	0.063	0.685	2842	<0.001	4.977	-0.263	0.061	0.699	2841	<0.001	
New York	genus	1.220	0.007	0.878	4801	<0.001	1.305	-0.313	0.006	0.922	4800	<0.001	*
	family	2.783	0.028	0.678	4801	<0.001	3.040	-0.321	0.027	0.723	4800	<0.001	*
	order	4.132	0.053	0.563	4801	<0.001	4.455	-0.260	0.054	0.593	4800	<0.001	*
Pennsylvania	genus	1.270	0.003	0.973	4786	<0.001	1.271	0.046	0.003	0.974	4785	<0.001	
	family	3.231	0.024	0.786	4786	<0.001	3.249	0.119	0.024	0.791	4785	<0.001	
	order	6.193	0.075	0.591	4786	<0.001	6.209	-0.056	0.075	0.592	4785	<0.001	
Washington	genus	1.205	0.005	0.980	1203	<0.001	1.217	-0.048	0.005	0.981	1202	<0.001	
	family	2.332	0.023	0.896	1203	<0.001	2.358	-0.050	0.024	0.897	1202	<0.001	
	order	4.225	0.100	0.597	1203	<0.001	4.158	0.063	0.104	0.599	1202	<0.001	

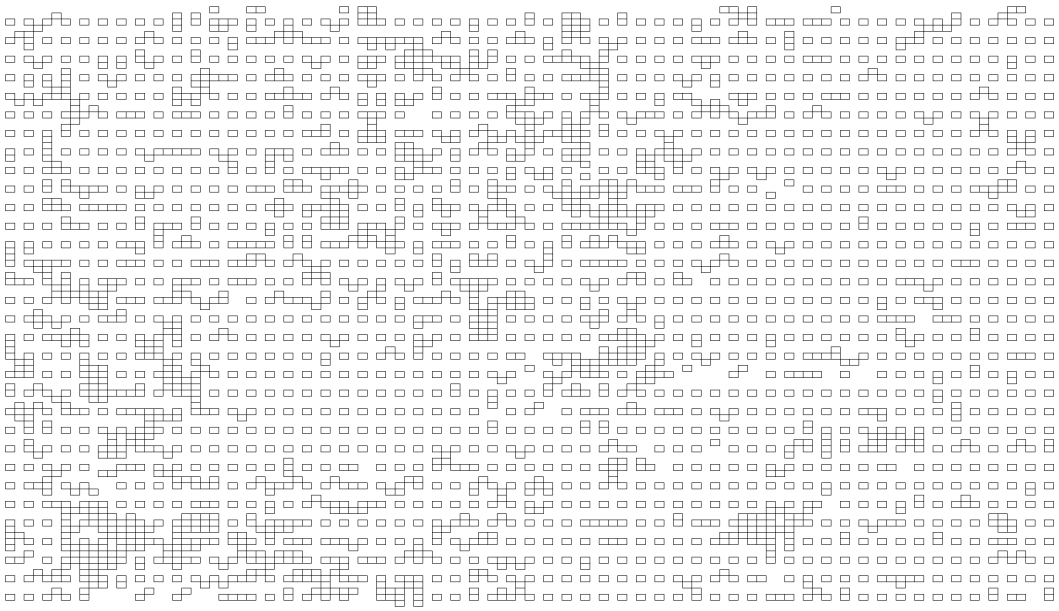
[a]Higher Taxon Coefficient

[b]Human Influence Coefficient

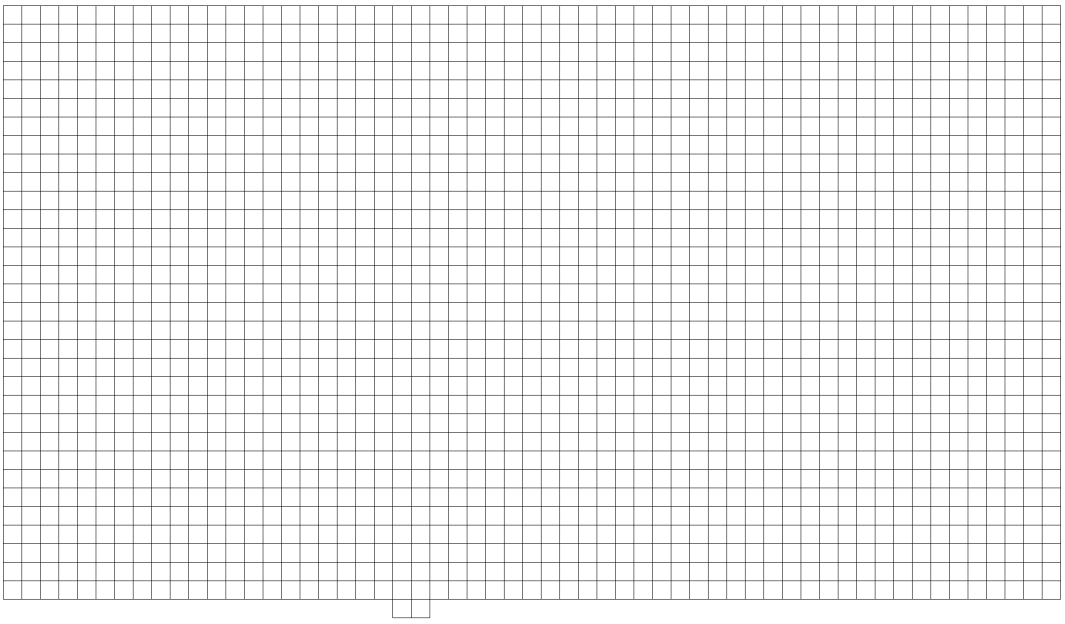
*denotes significant difference between models with human influence and models without human influence

Figure 1. Maps of sample states broken into blocks and quads: Colorado Breeding Bird Atlas (BBA) blocks (a) and quads (b); Florida BBA blocks (c) and quads (d); Michigan BBA blocks (e) and quads (f); New York BBA blocks (g) and (h); Pennsylvania BBA blocks (i) and quads (j); Washington BBA blocks (k) and quads (l).

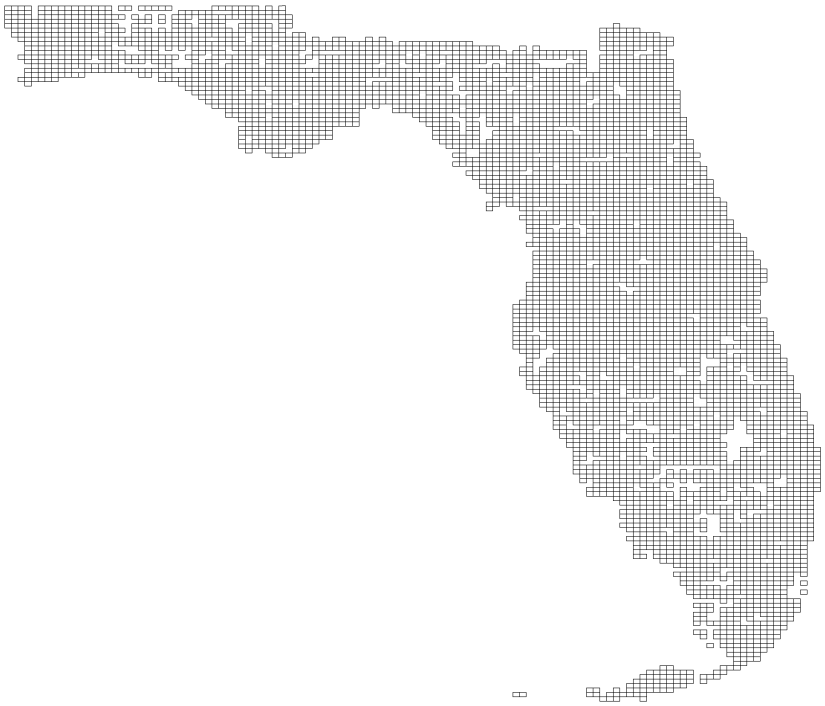
(a)



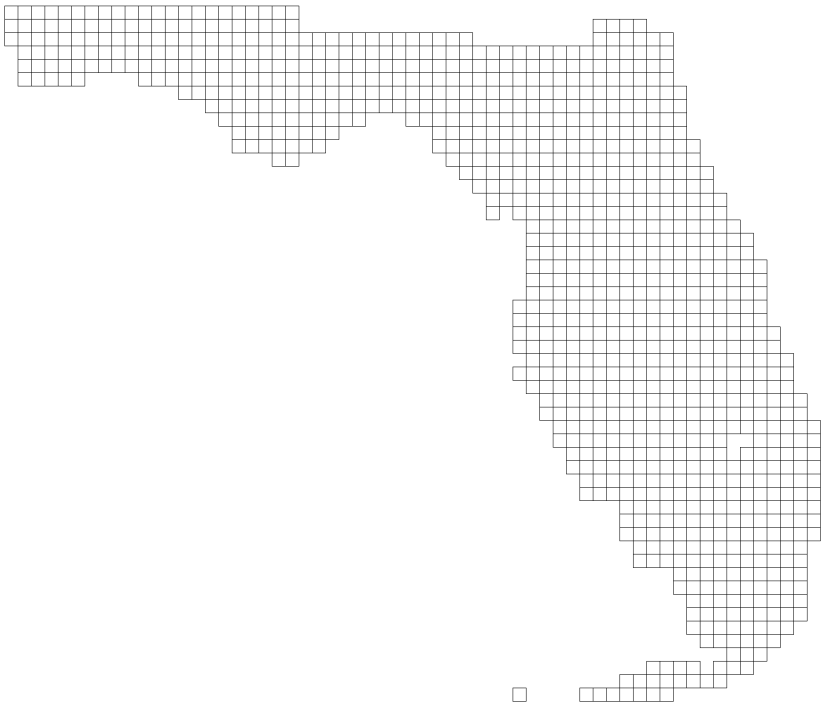
(b)



(c)



(d)

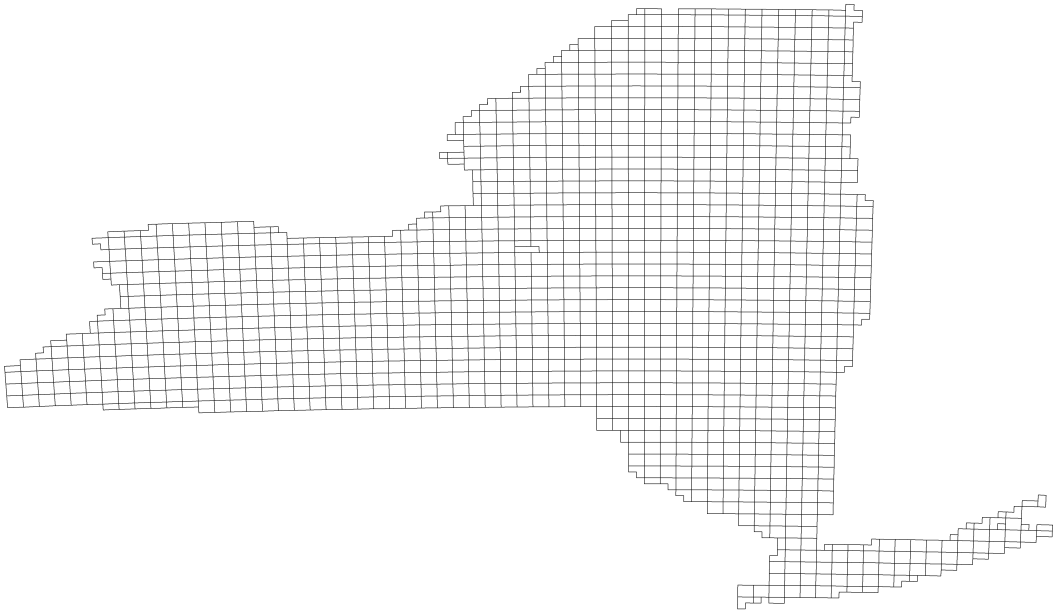


(e)

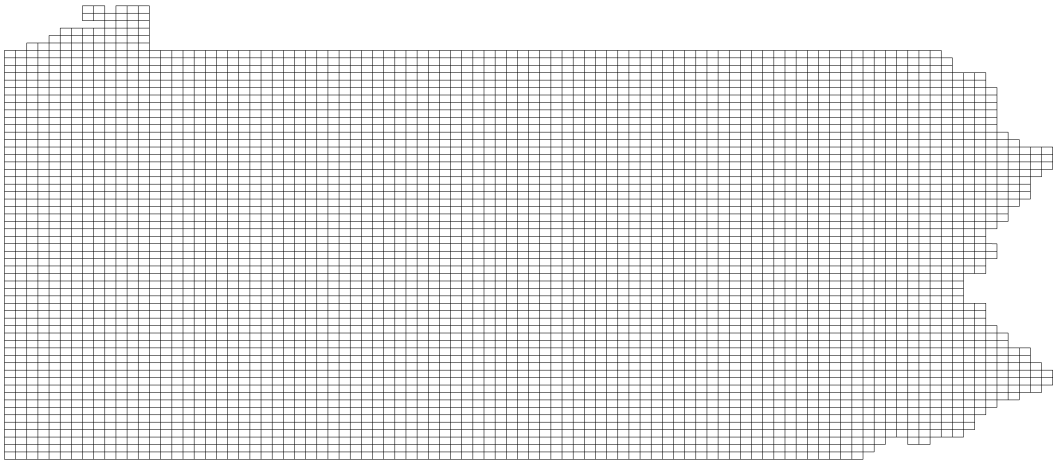
(f)

(g)

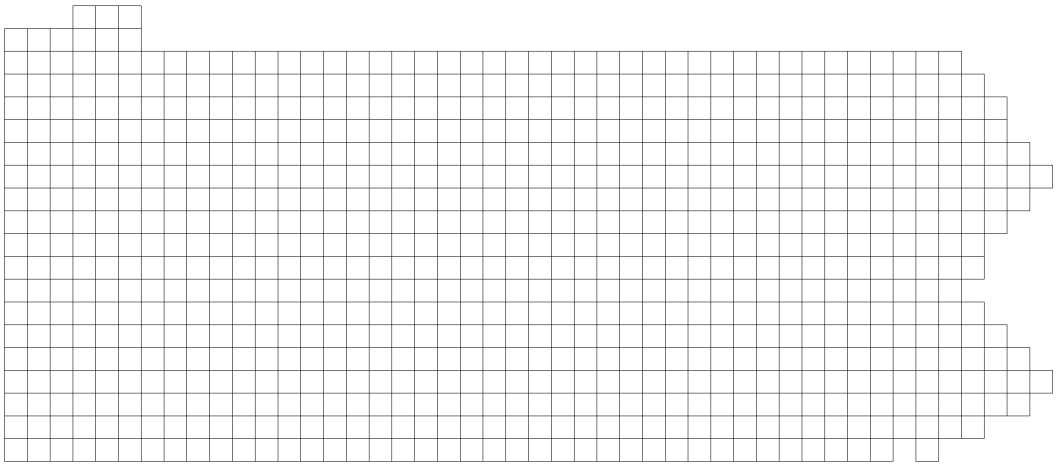
(h)



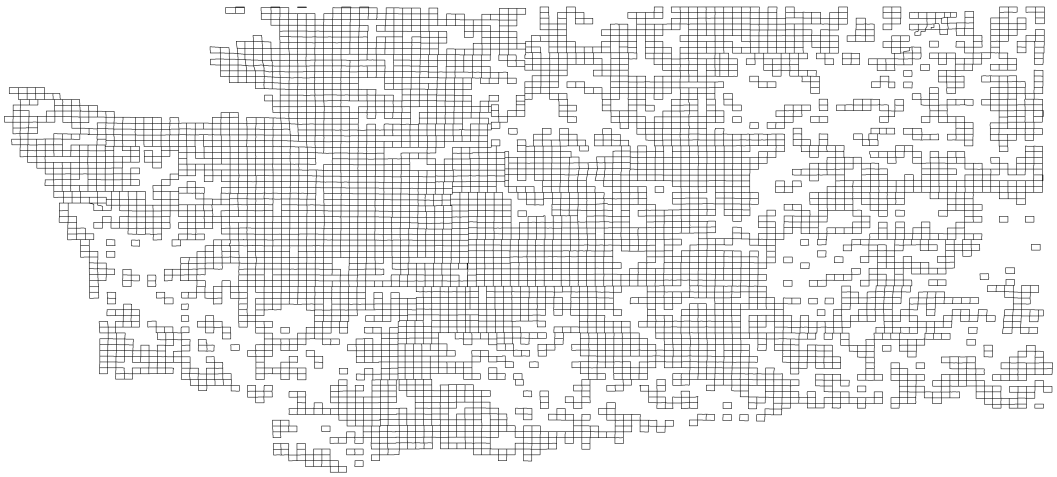
(i)



(j)



(k)



(l)

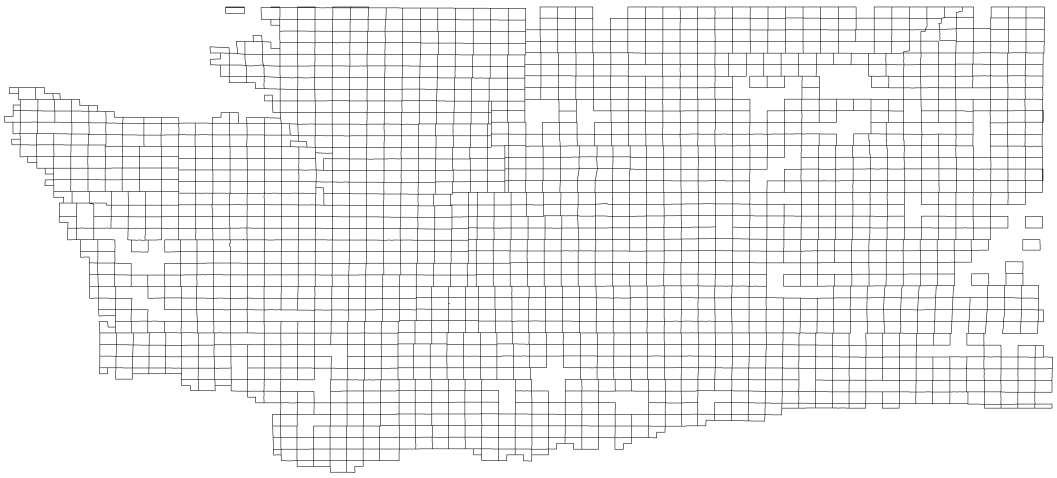
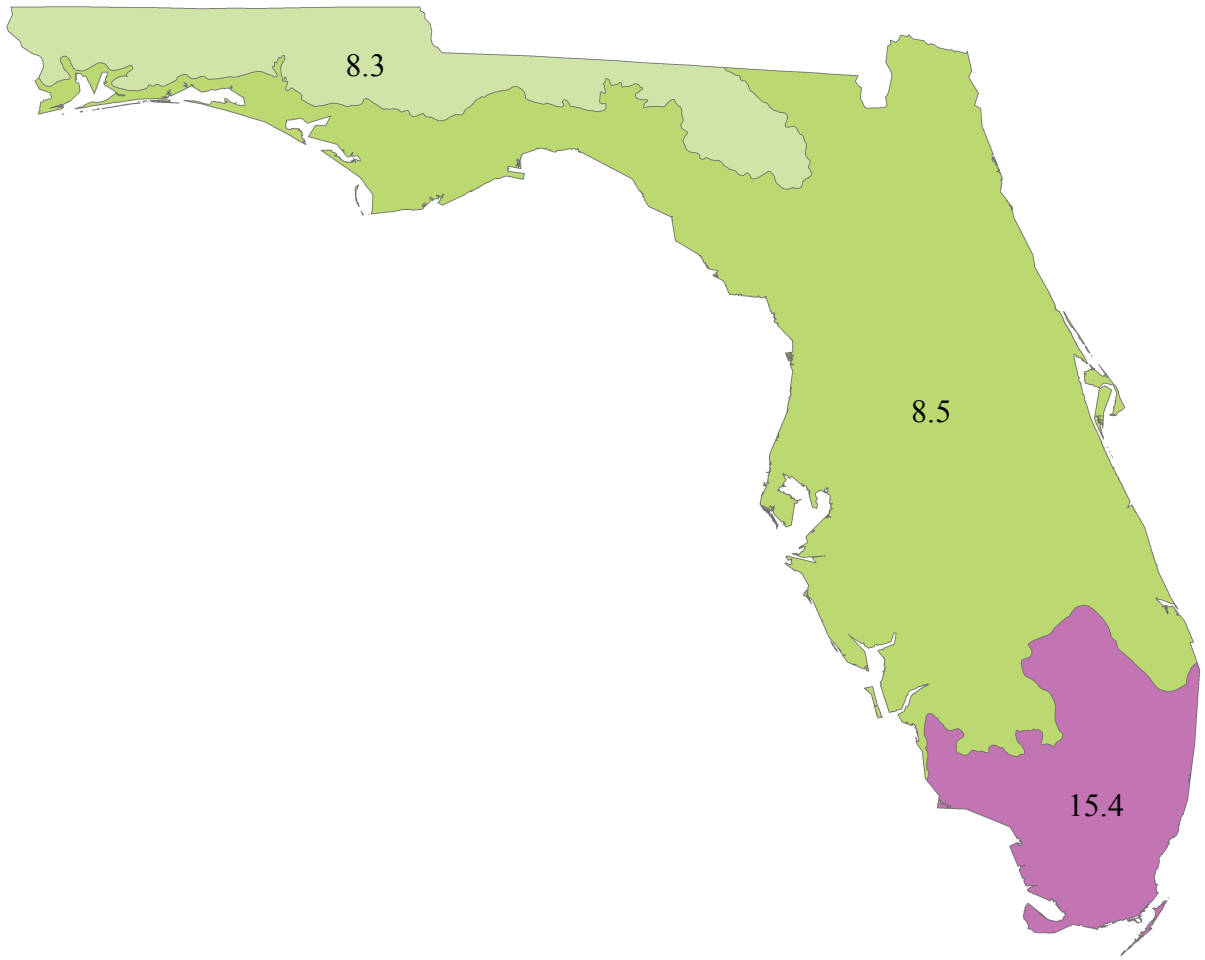


Figure 2. Maps of the ecoregions used for analysis within each state: Colorado (a); Florida (b); Michigan (c); New York (d); Pennsylvania (e); Washington (f).

(a)

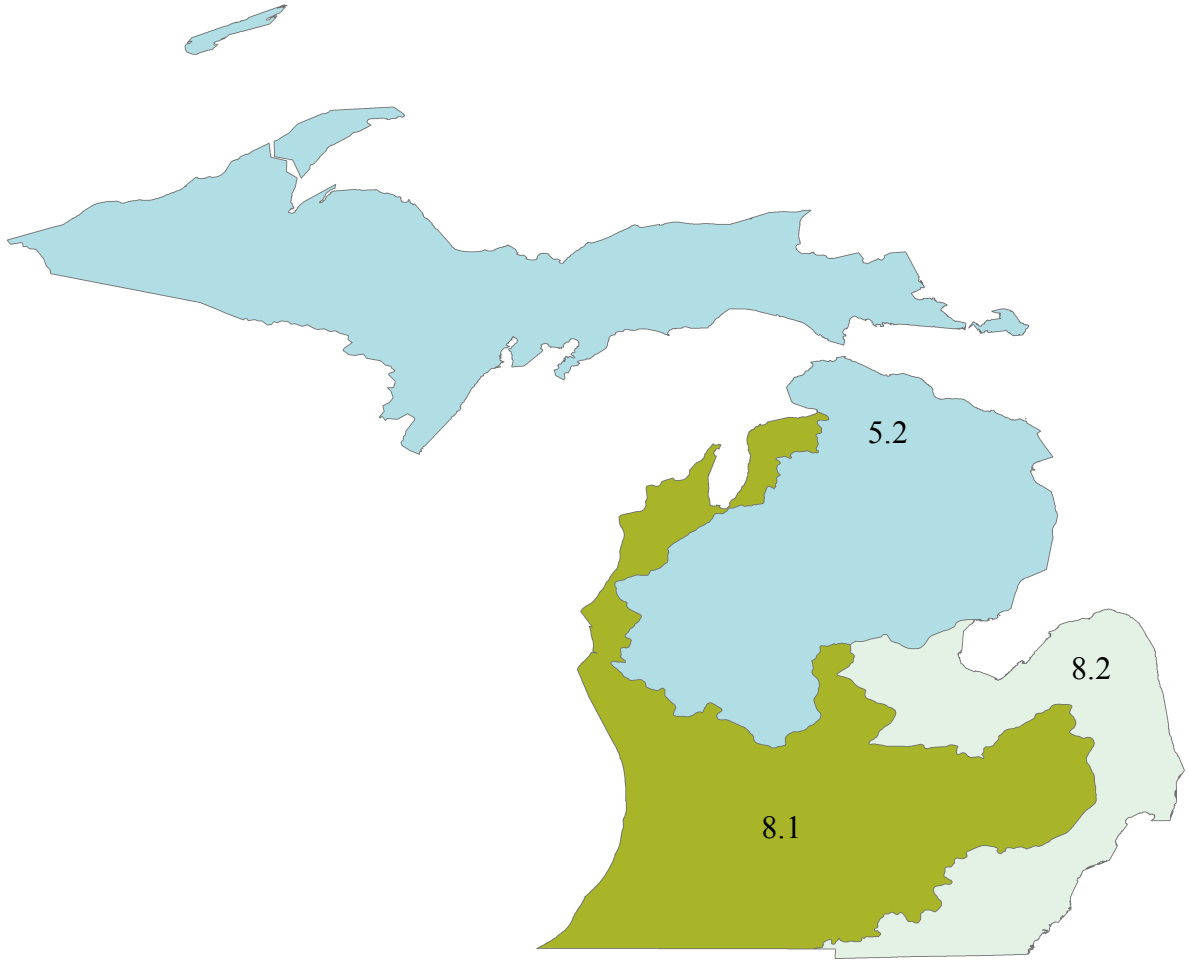
Colorado ecoregions: 6.2 Western Cordillera, 9.4 South Central Semi-arid Prairies, 10.1 Cold Deserts.

(b)



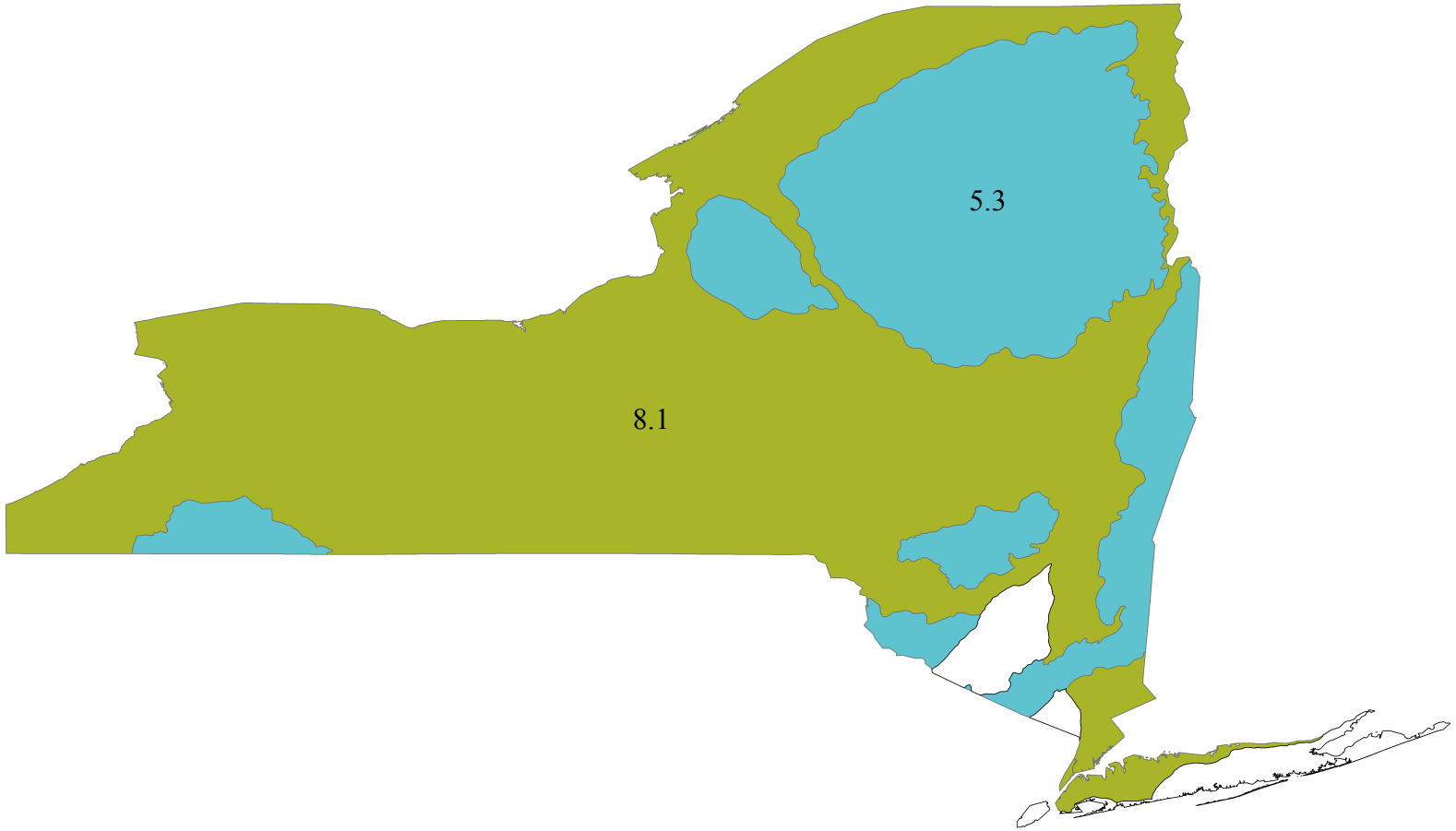
Florida ecoregions: 8.3 Southeastern USA Plains, 8.5 Mississippi Alluvial and Southeast USA Coastal Plains, 15.4 Everglades.

(c)



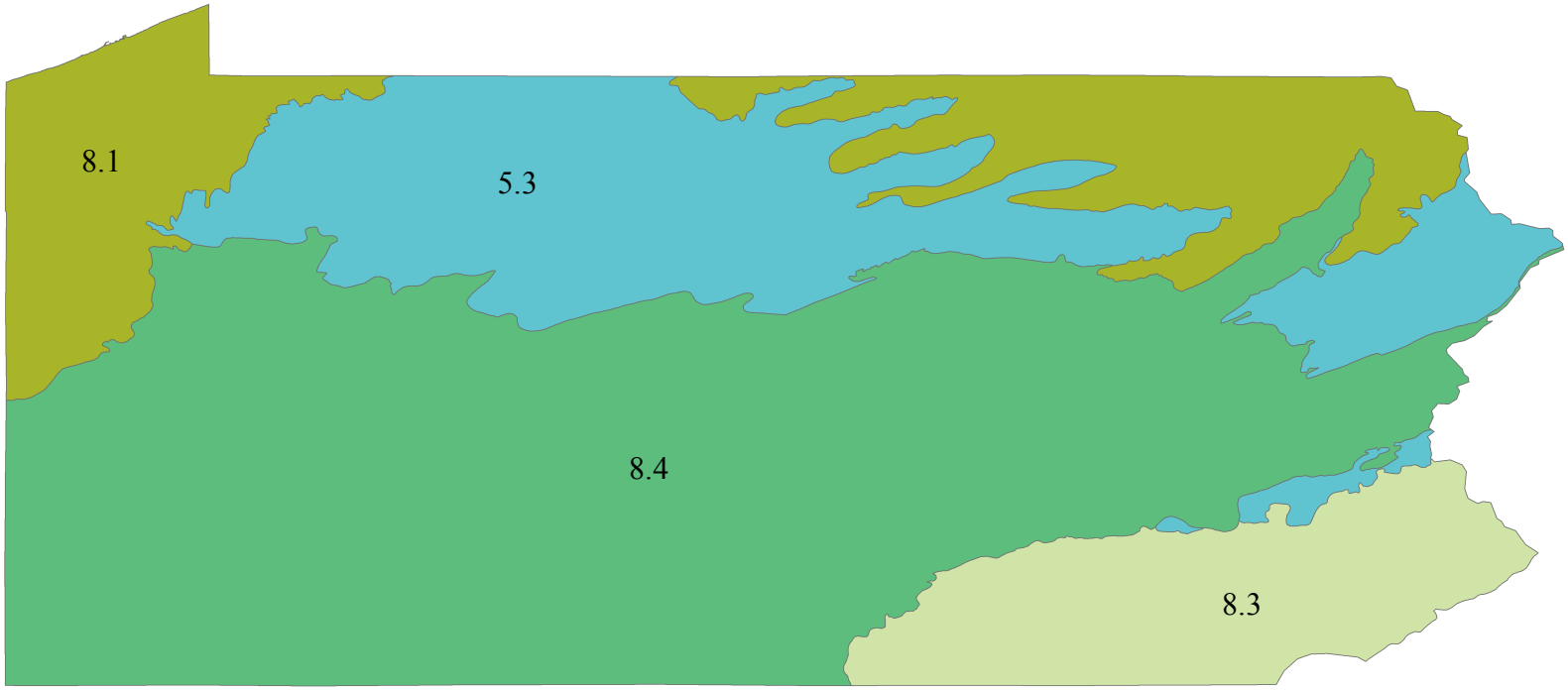
Michigan ecoregions: 5.2 Mixed Wood Shield, 8.1 Mixed Wood Plains, 8.2 Central USA Plains.

(d)



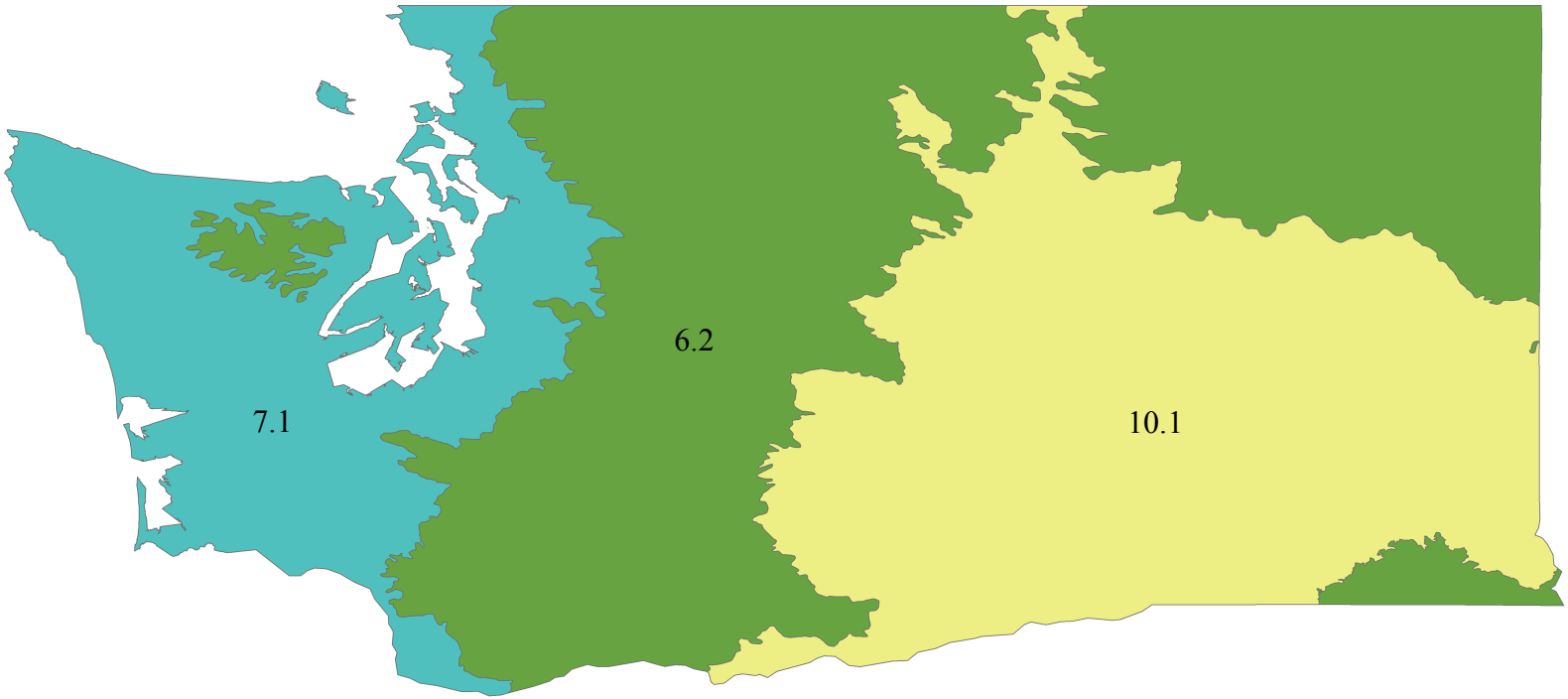
New York ecoregions: 5.3 Atlantic Highlands, 8.1 Mixed Wood Plains

(e)



Pennsylvania ecoregions: 5.3 Atlantic Highlands, 8.1 Mixed Wood Plains, 8.3 Southeastern USA Plains, 8.4 Ozark, Ouachita-Appalachian Forests.

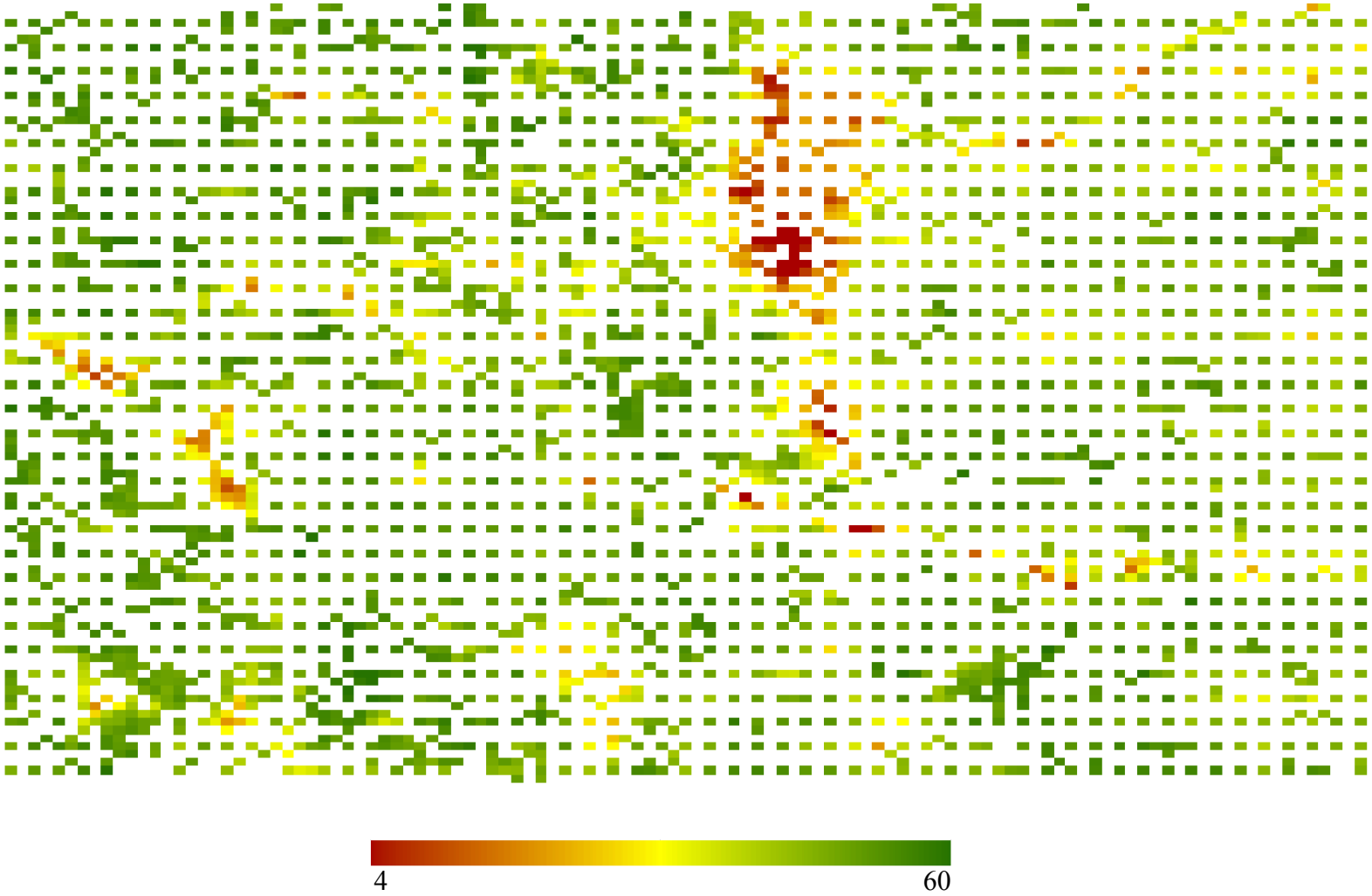
(f)



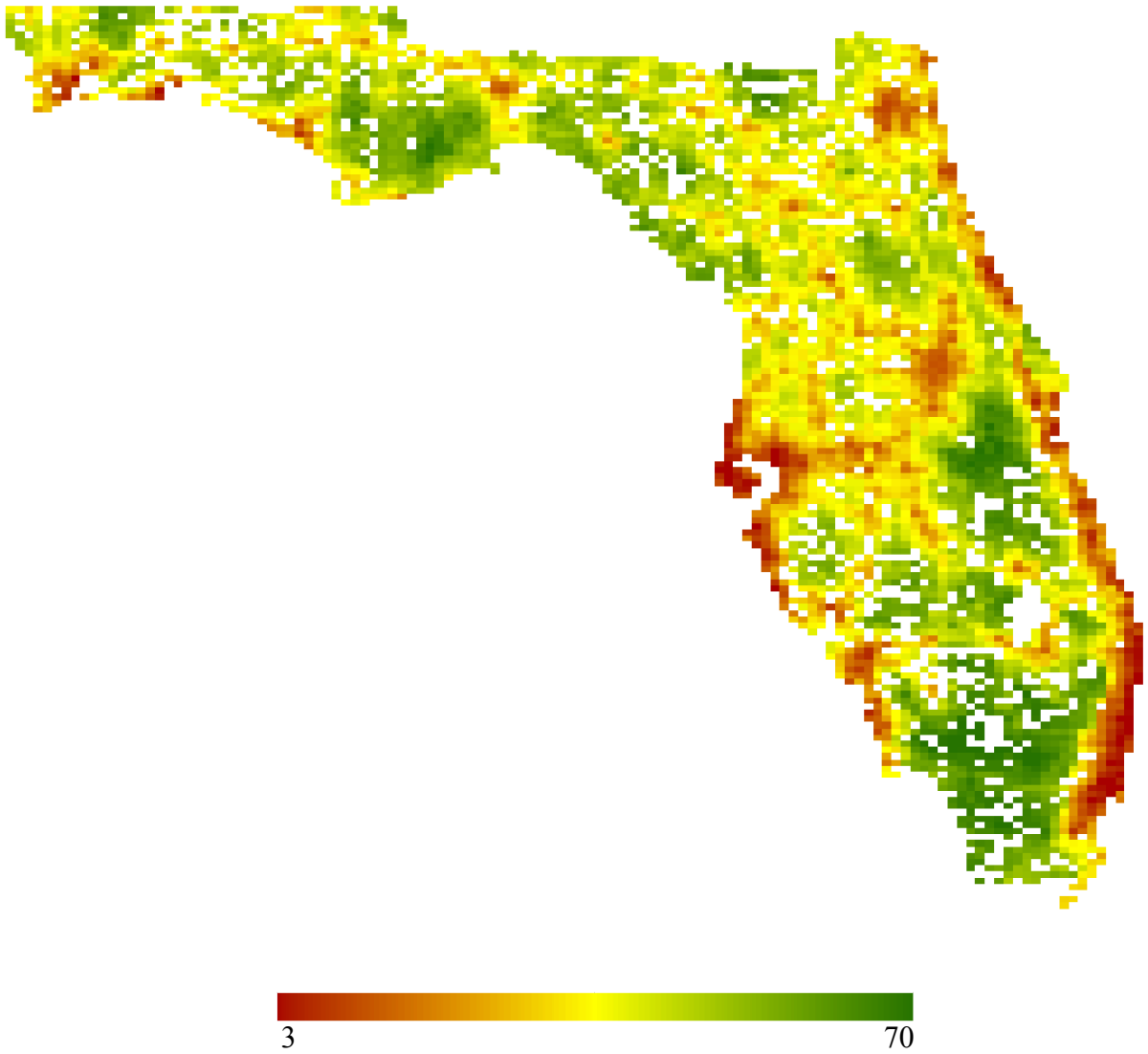
Washington ecoregions: 6.2 Western Cordillera, 7.1 Marine West Coast Forest, 10.1 Cold Deserts.

Figure 3. Maps of human influence values in sample states, broken into blocks: Colorado Breeding Bird Atlas (BBA) (a); Florida BBA (b); Michigan BBA (c); New York BBA (d); Pennsylvania BBA (e); Washington BBA (f)

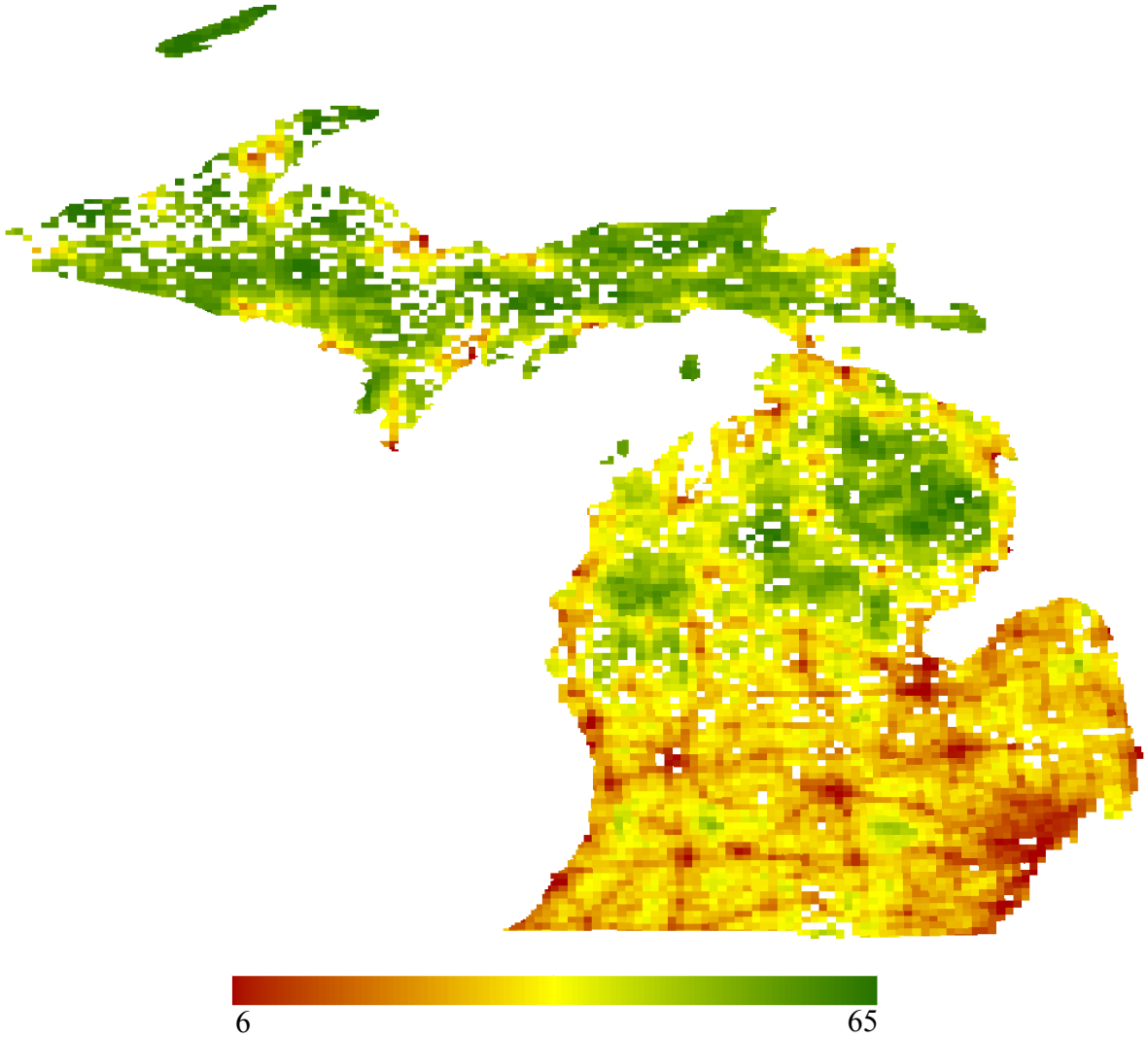
(a)



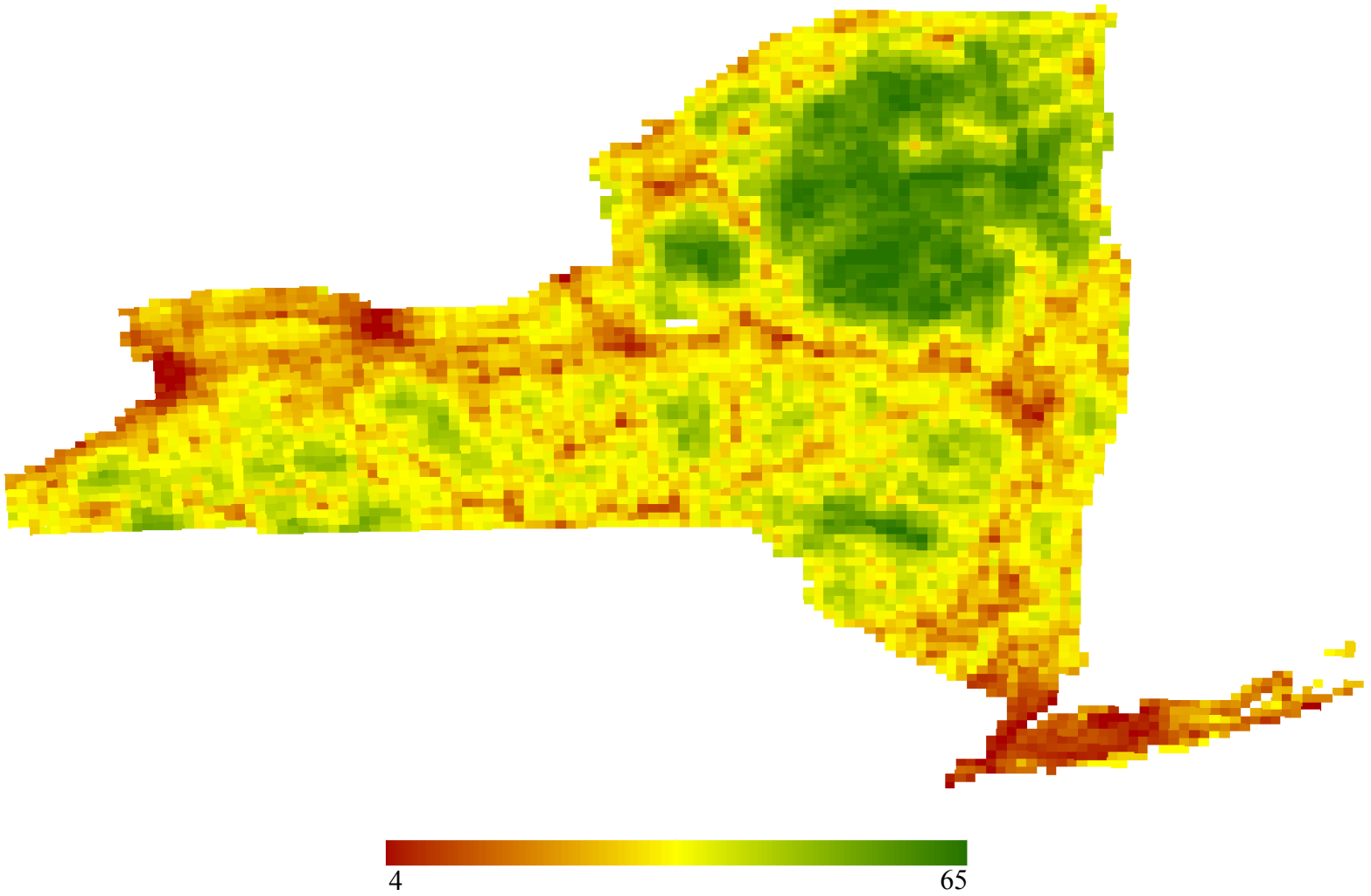
(b)



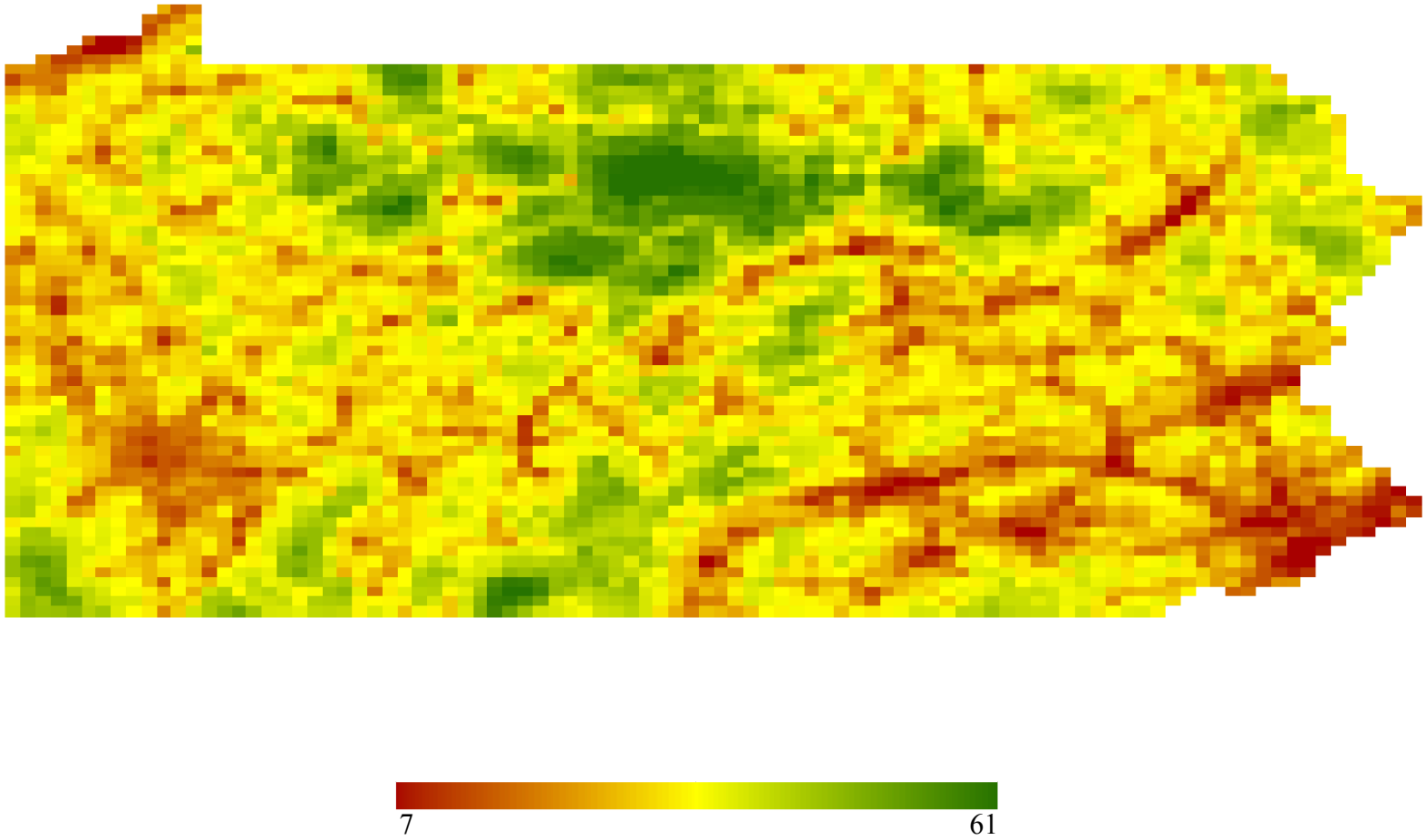
(c)



(d)



(e)



(f)

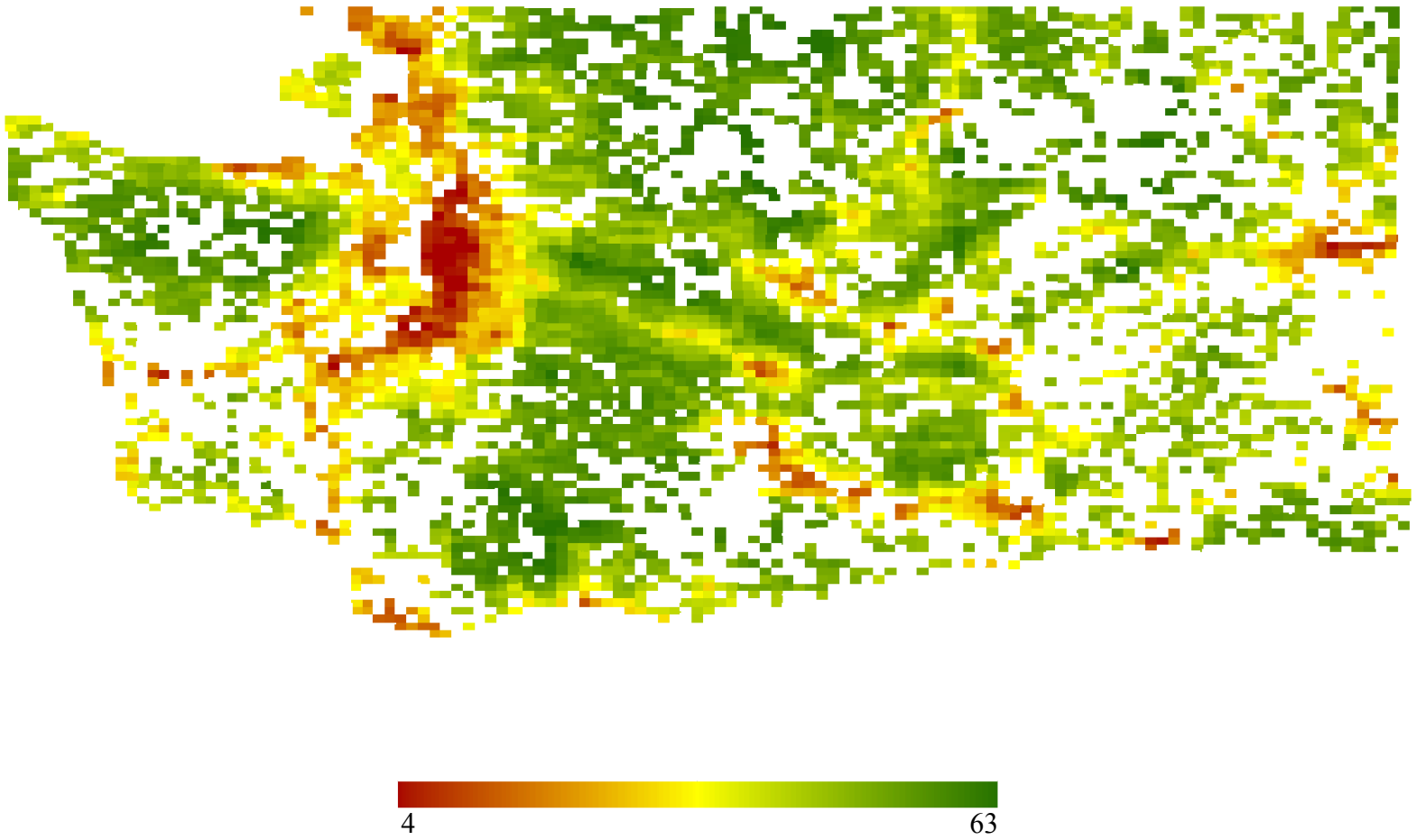
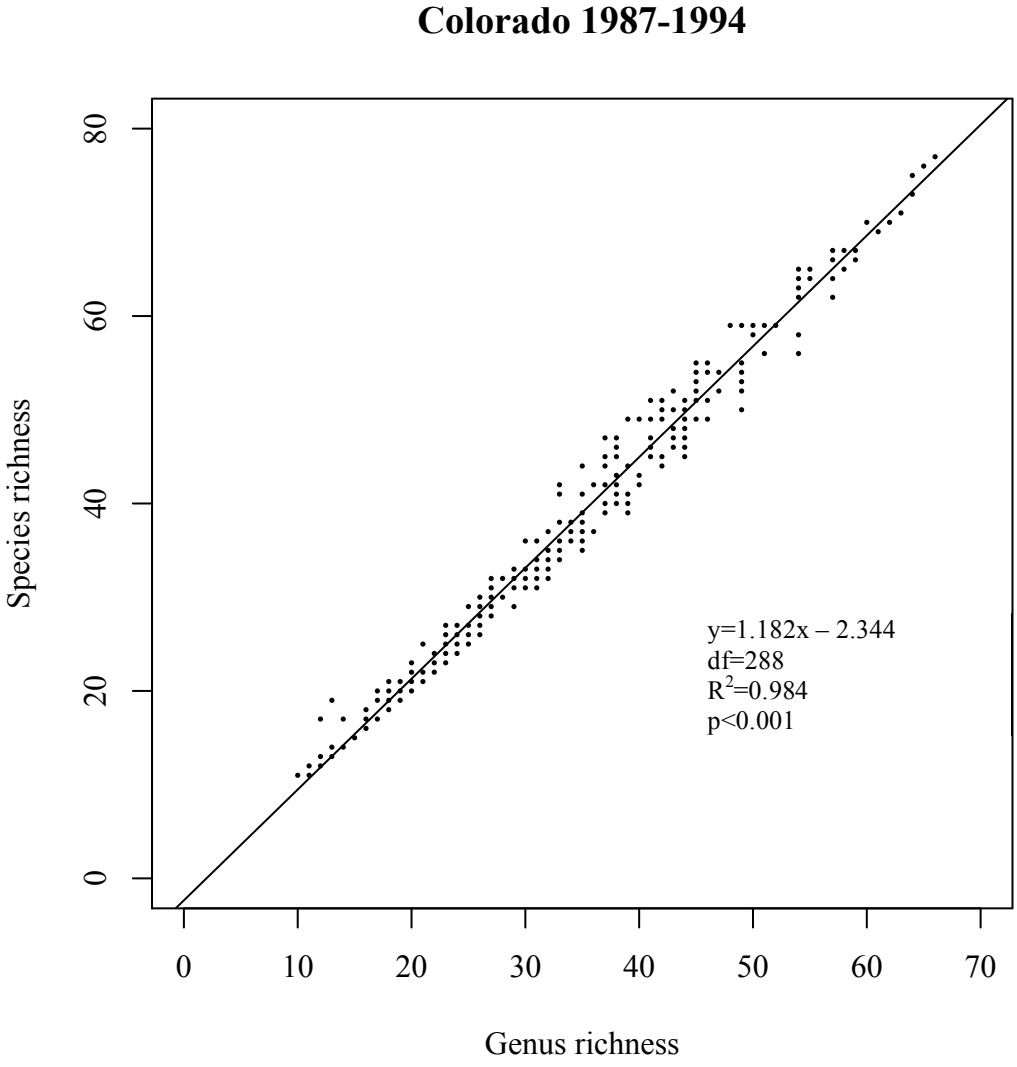
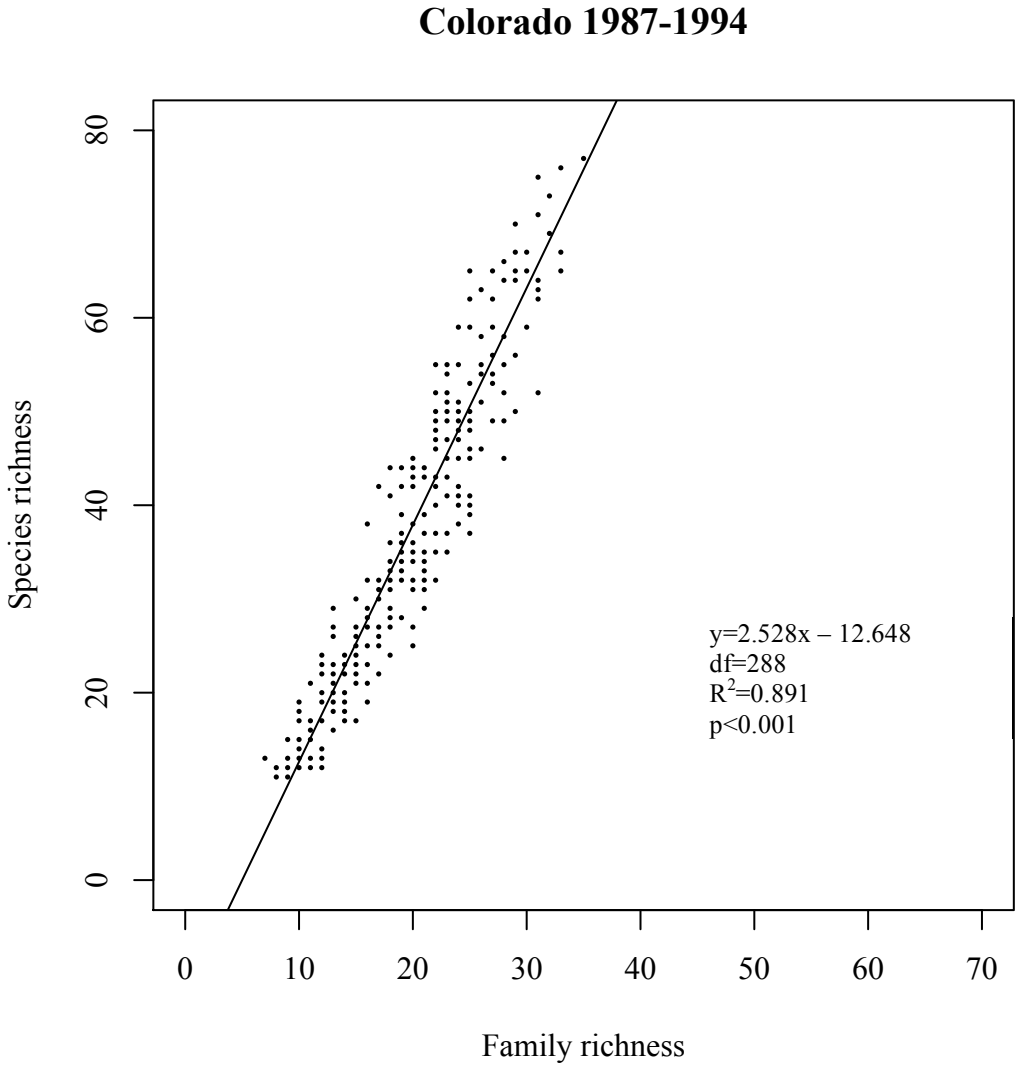


Figure 4. Scatterplot graphs illustrating regressions analysis for the basic surrogacy models in all six states: Colorado genus richness (a), family richness (b), and order richness (c); Florida genus richness (a), family richness (b), and order richness (c); Michigan genus richness (a), family richness (b), and order richness (c); New York genus richness (a), family richness (b), and order richness (c); Pennsylvania genus richness (a), family richness (b), and order richness (c); and Washington genus richness (a), family richness (b), and order richness (c).

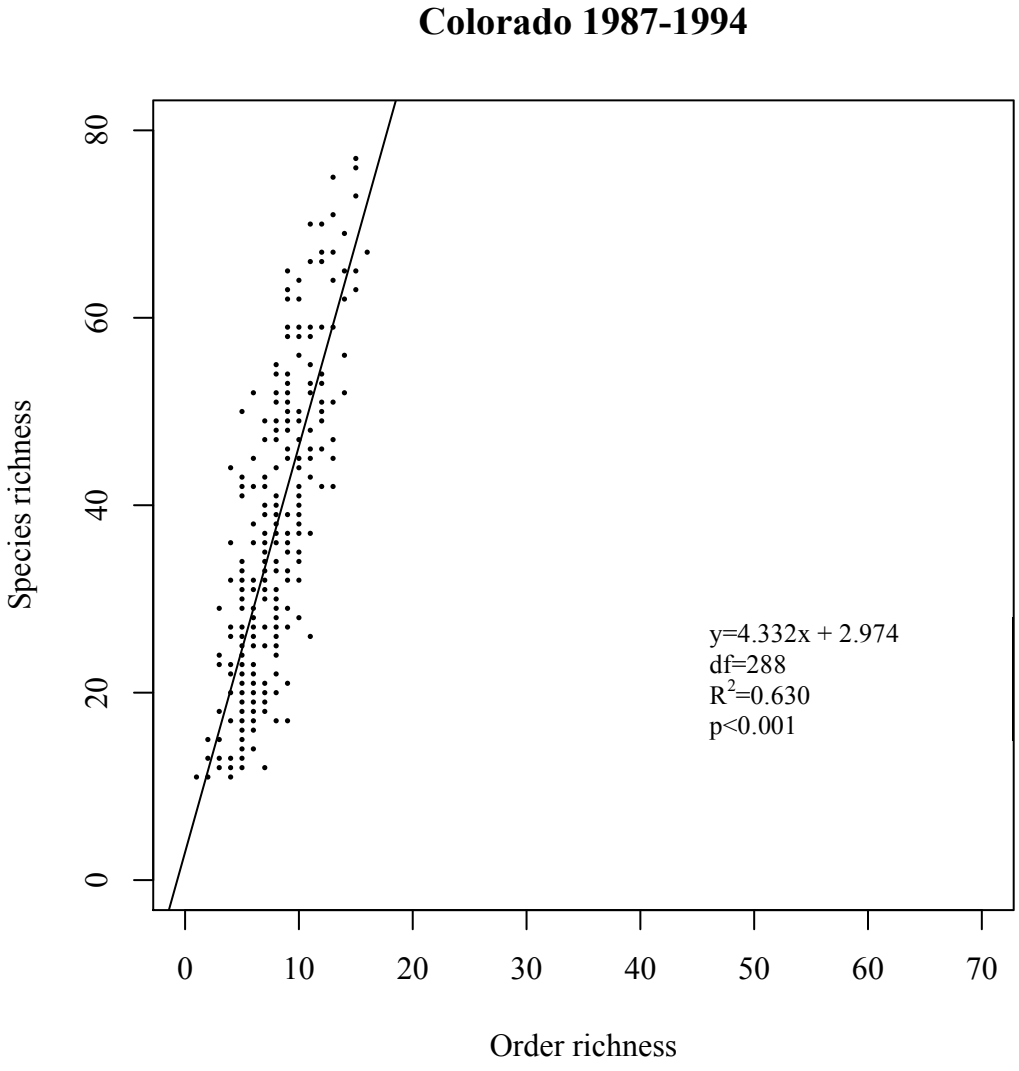
(a)


Colorado 1987-1994
Species richness
Genus richness
0
20
40
60
80
0
10
20
30
40
50
60
70
y=1.182x – 2.344
df=288
R2=0.984
p<0.001

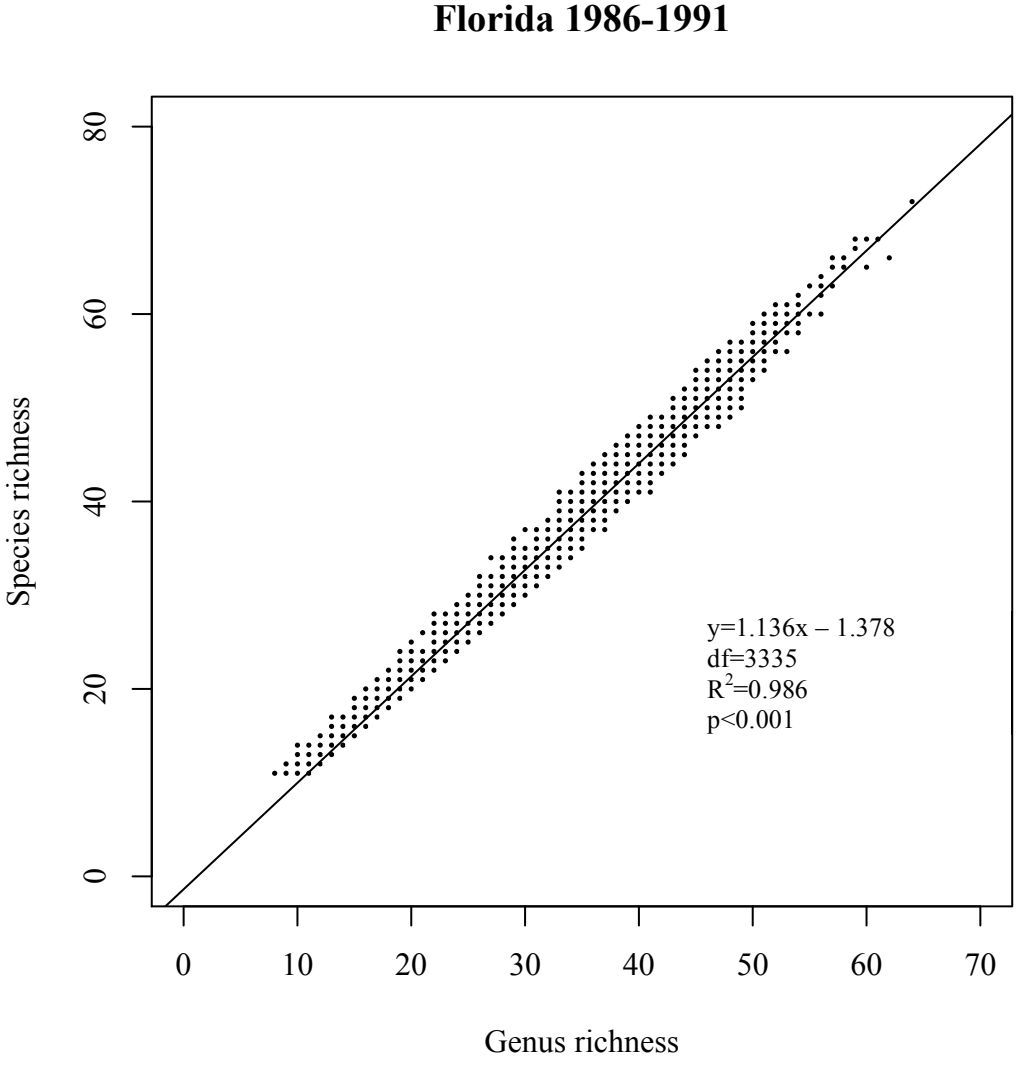
(b)


Colorado 1987-1994
Species richness
Family richness
0
20
40
60
80
0
10
20
30
40
50
60
70
y=2.528x – 12.648
df=288
R2=0.891
p<0.001

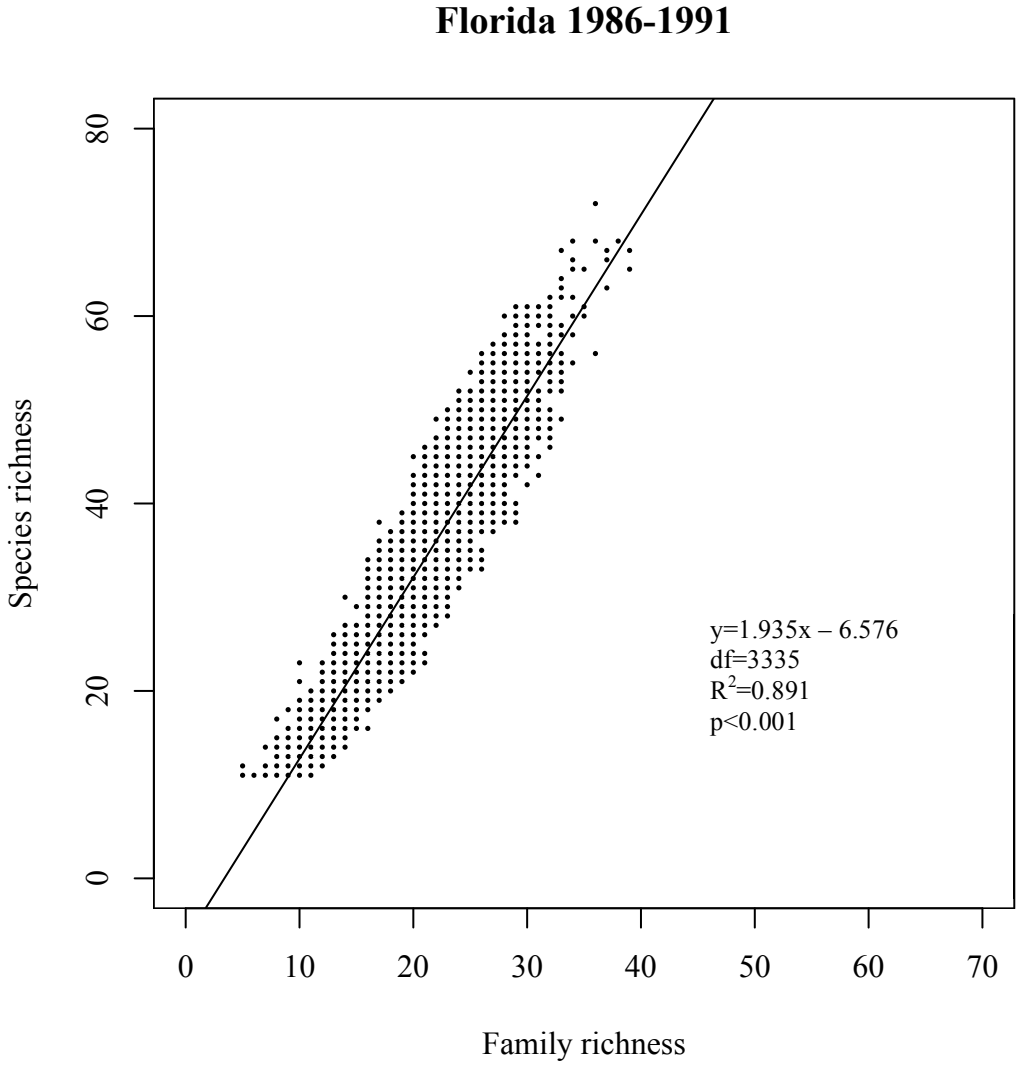
(c)


Colorado 1987-1994
Species richness
Order richness
0
20
40
60
80
0
10
20
30
40
50
60
70
y=4.332x + 2.974
df=288
R2=0.630
p<0.001

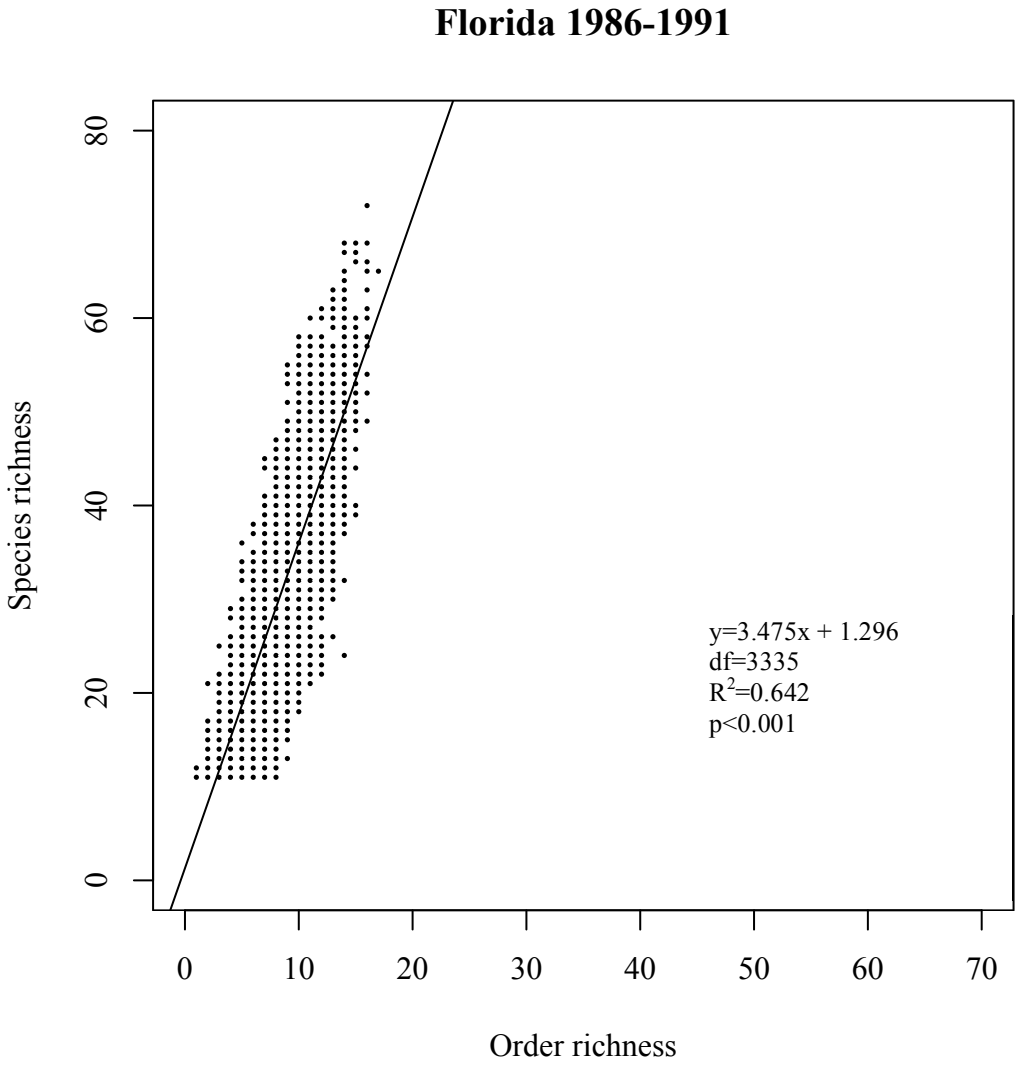
(a)


Florida 1986-1991
Species richness
Genus richness
0
20
40
60
80
0
10
20
30
40
50
60
70
y=1.136x – 1.378
df=3335
R2=0.986
p<0.001

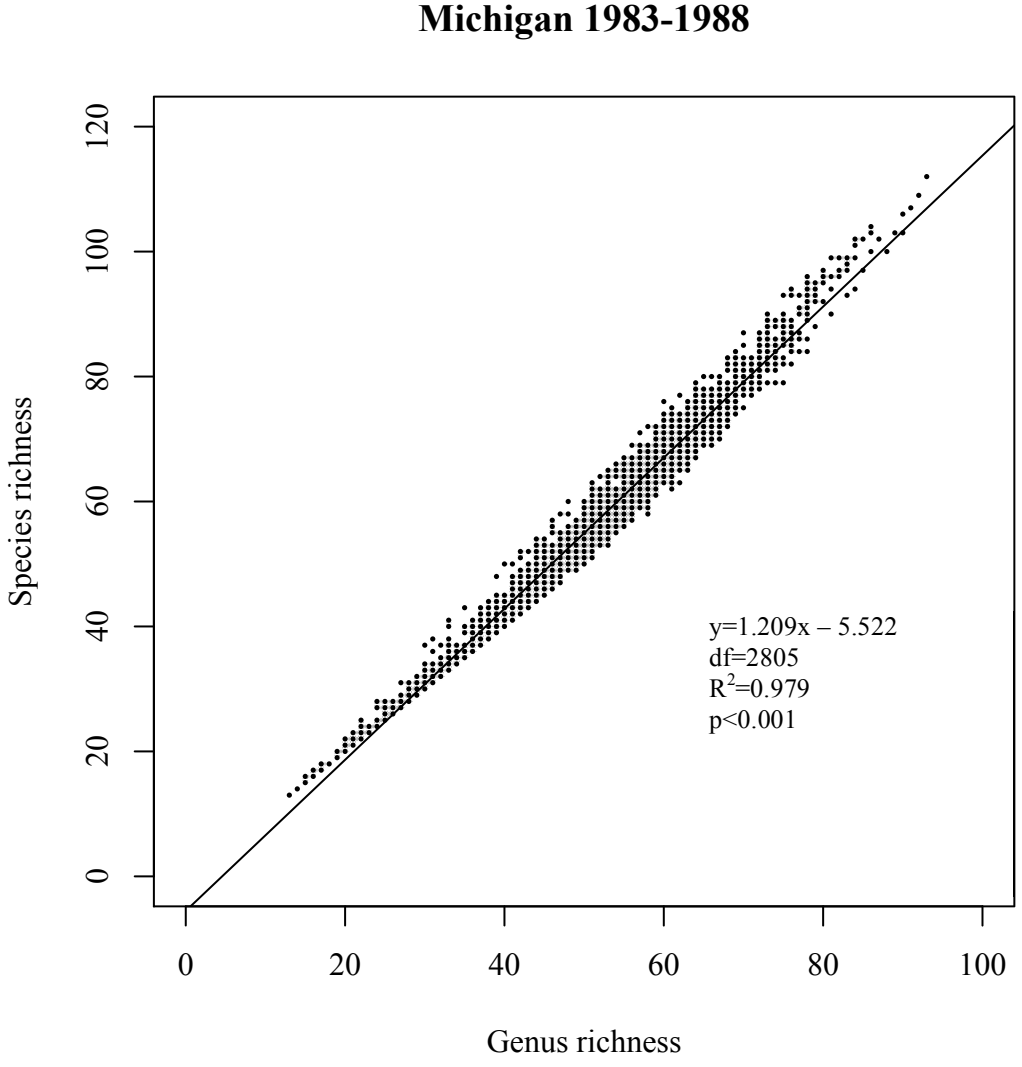
(b)


Florida 1986-1991
Species richness
Family richness
0
20
40
60
80
0
10
20
30
40
50
60
70
y=1.935x – 6.576
df=3335
R2=0.891
p<0.001

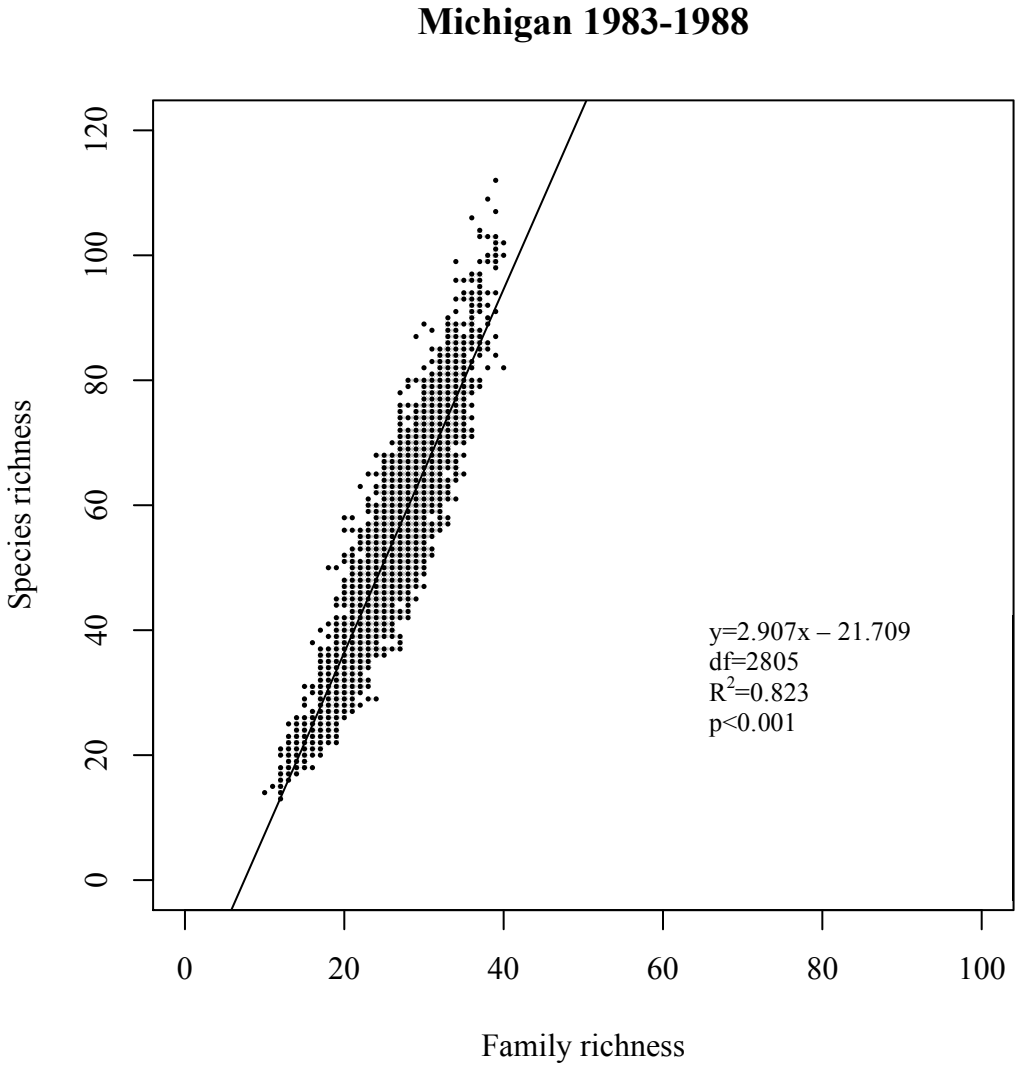
(c)


Florida 1986-1991
Species richness
0
20
40
60
80
0
10
20
30
40
50
60
70
Order richness
y=3.475x + 1.296
df=3335
R2=0.642
p<0.001

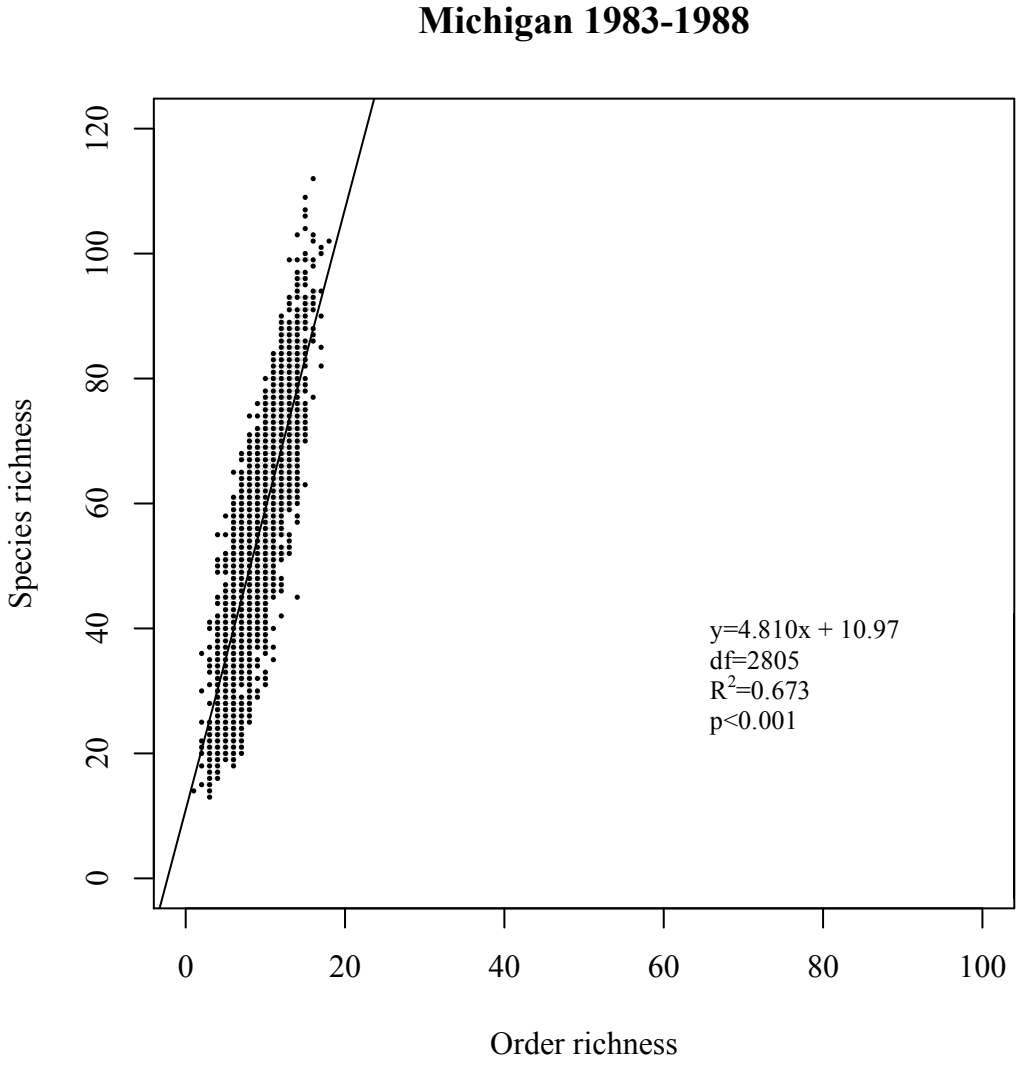
(a)


Michigan 1983-1988
Species richness
Genus richness
0
20
40
60
80
100
120
y=1.209x – 5.522
df=2805
R2=0.979
p<0.001

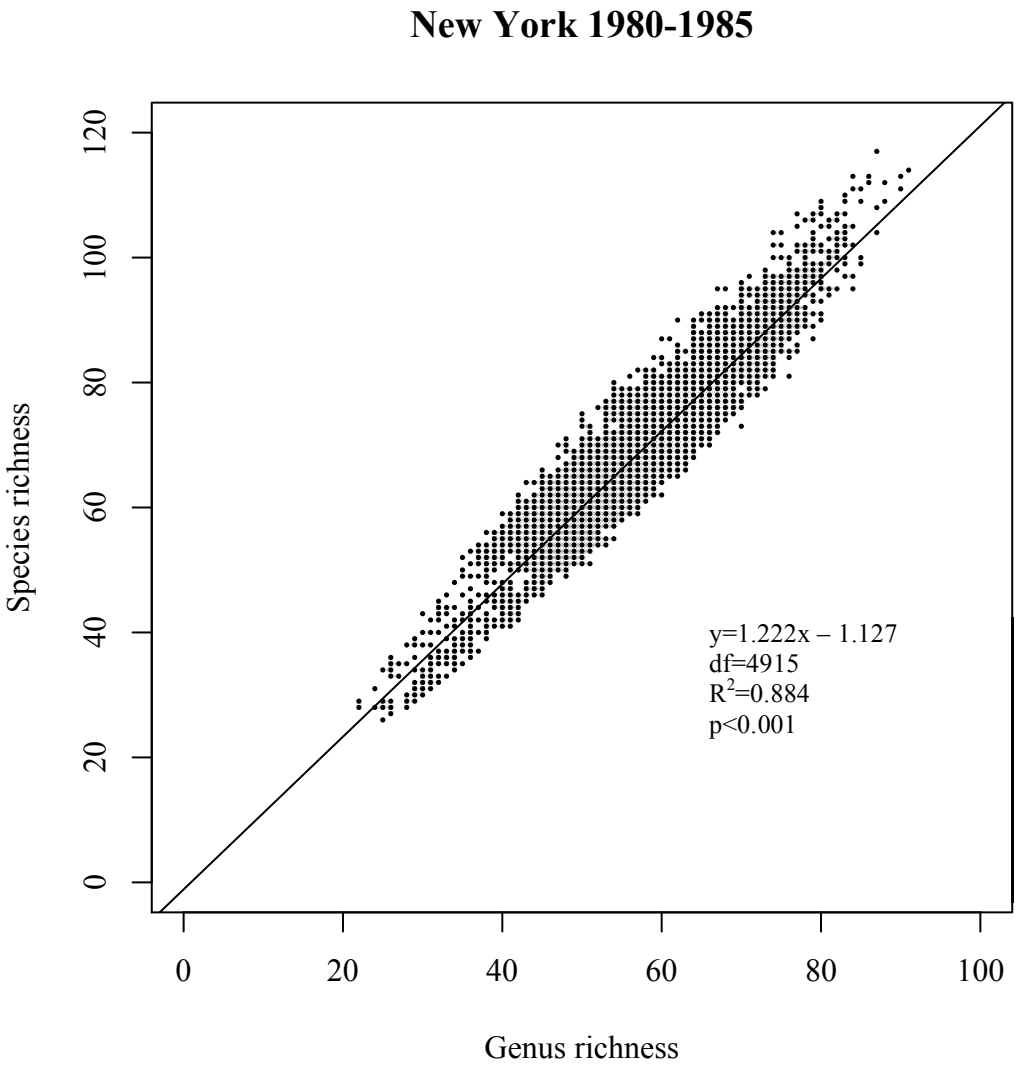
(b)


Michigan 1983-1988
Species richness
0
20
40
60
80
100
120
Family richness
0
20
40
60
80
100
y=2.907x – 21.709
df=2805
R2=0.823
p<0.001

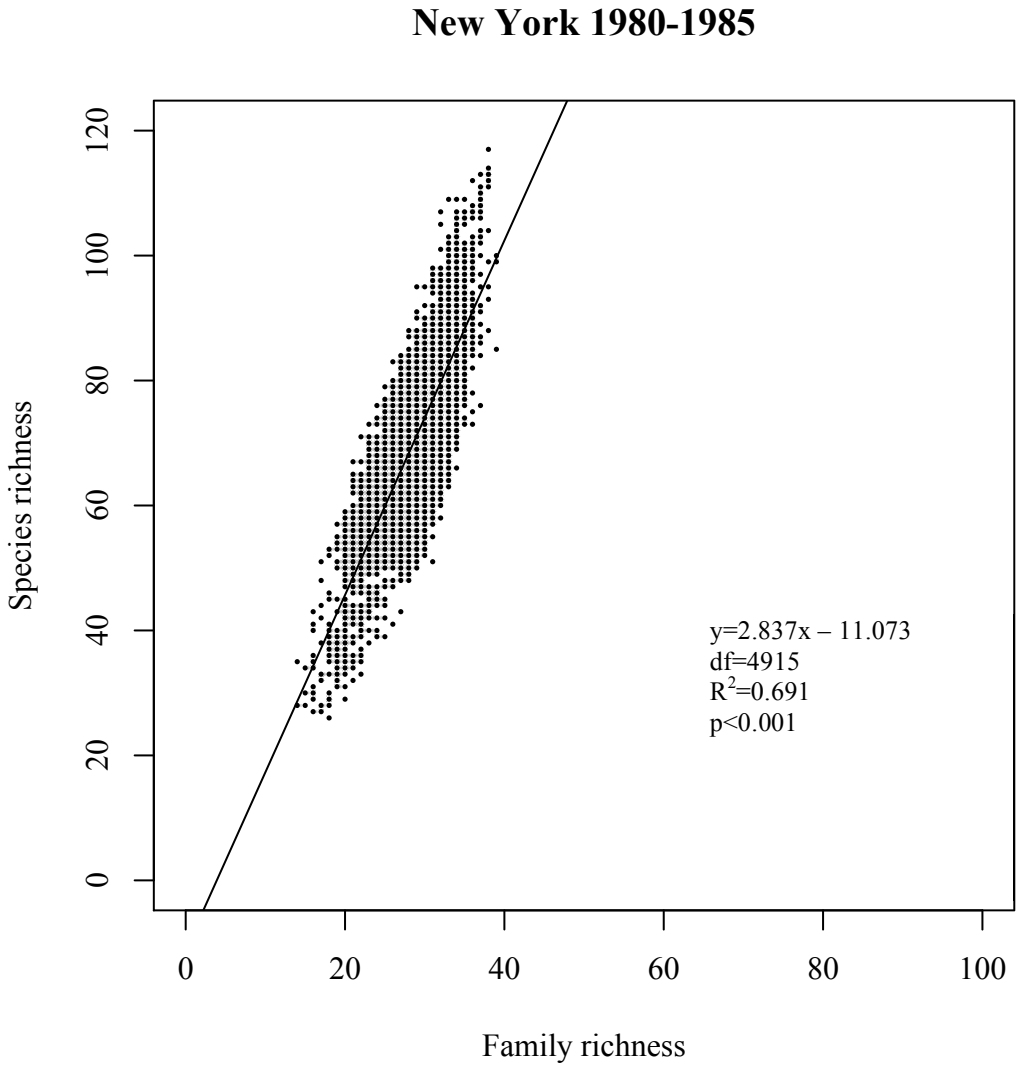
(c)


Michigan 1983-1988
Species richness
0
20
40
60
80
100
120
Order richness
0
20
40
60
80
100
y=4.810x + 10.97
df=2805
R2=0.673
p<0.001

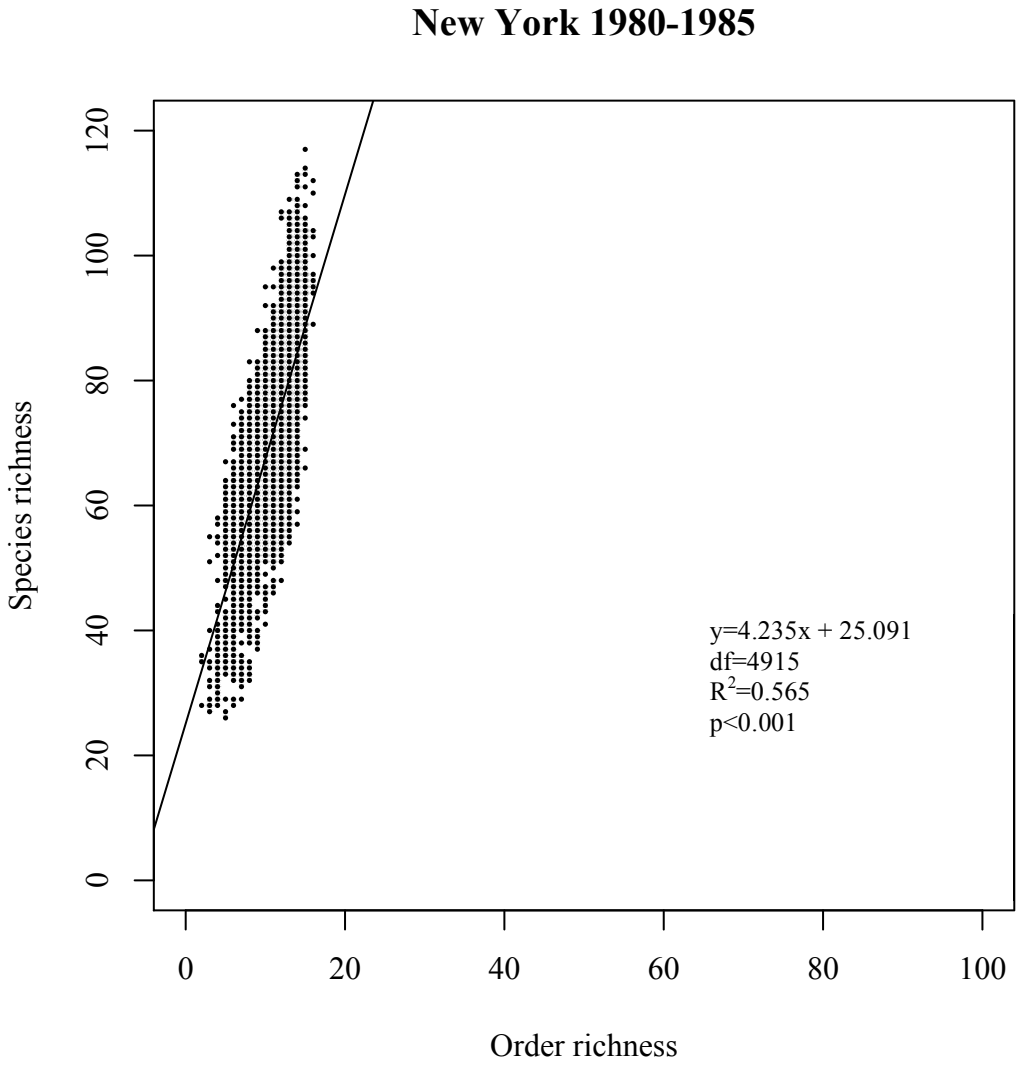
(a)


New York 1980-1985
Species richness
Genus richness
0
20
40
60
80
100
120
y=1.222x – 1.127
df=4915
R2=0.884
p<0.001

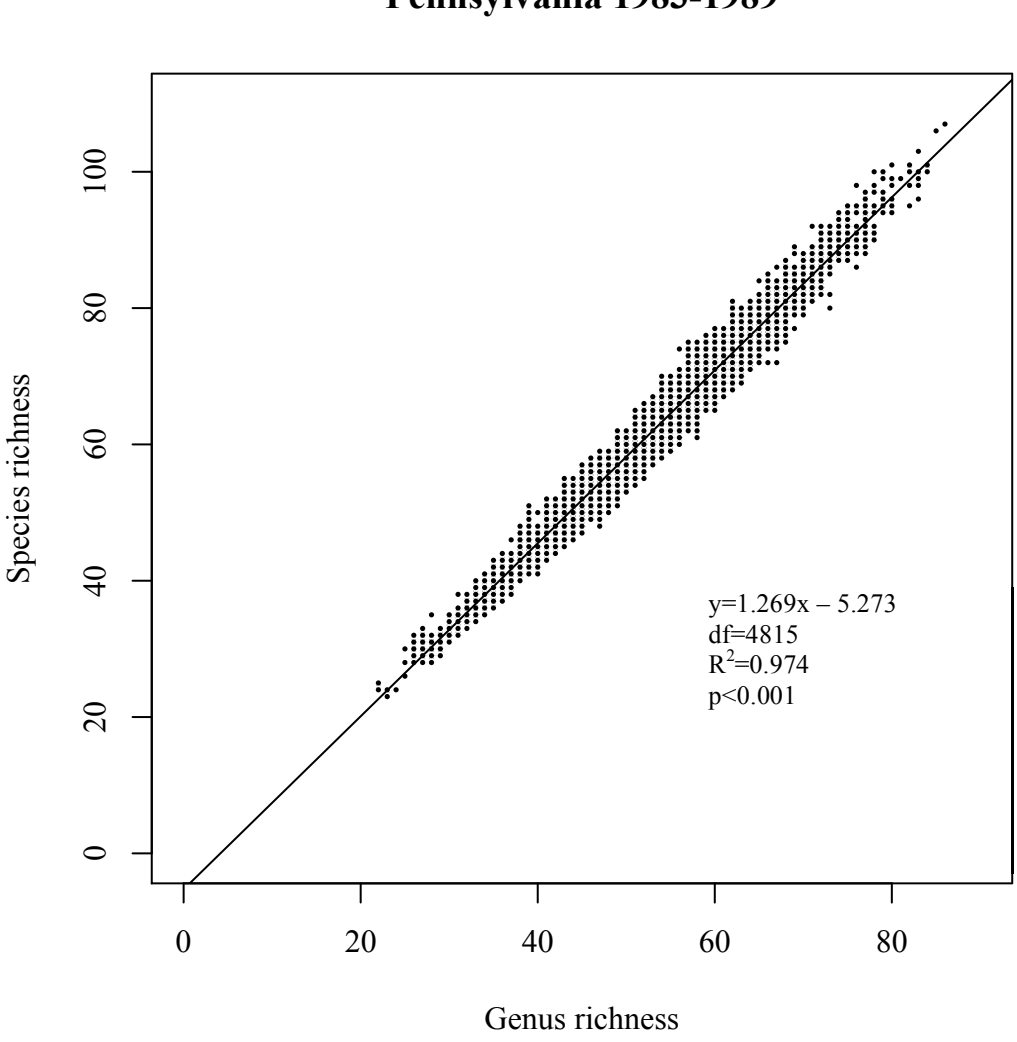
(b)


New York 1980-1985
Species richness
Family richness
0
20
40
60
80
100
120
y=2.837x – 11.073
df=4915
R2=0.691
p<0.001

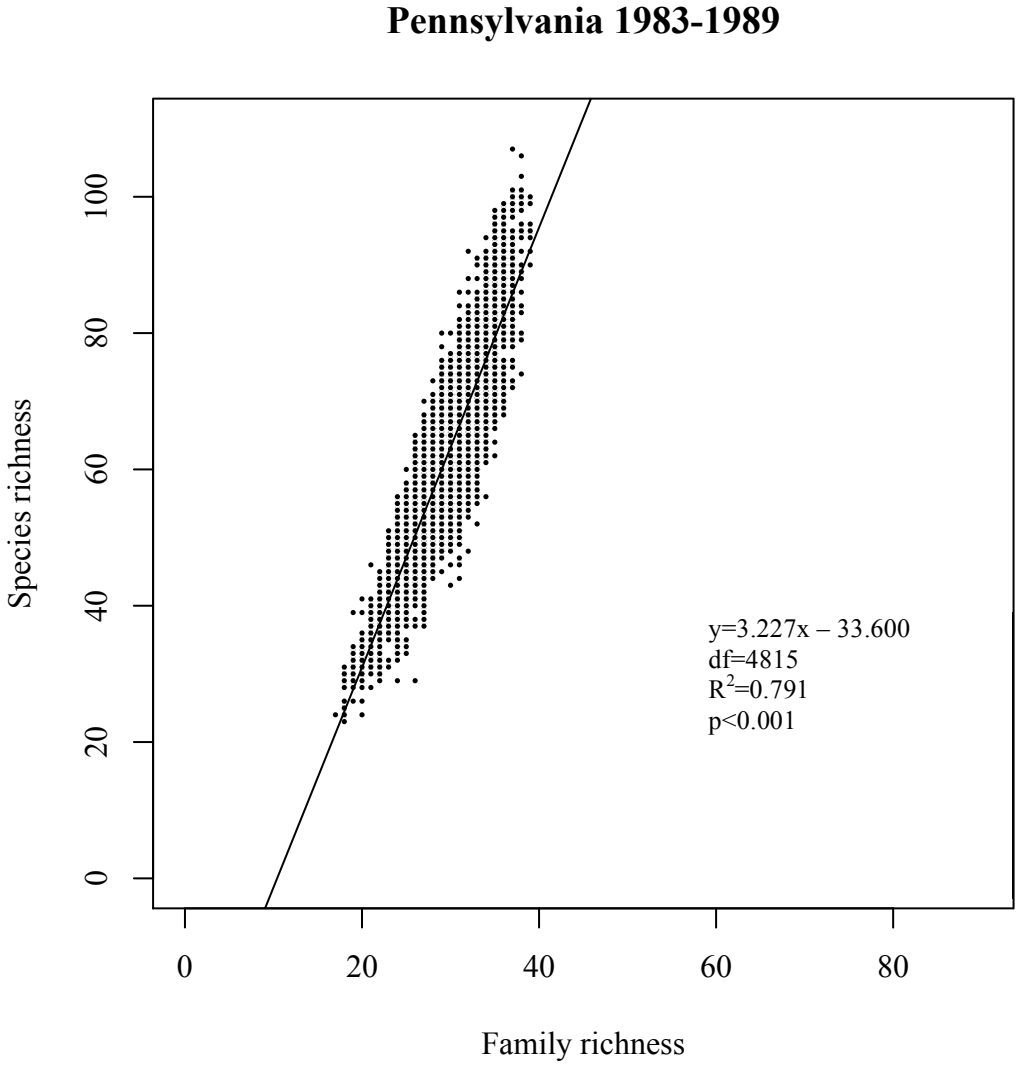
(c)


New York 1980-1985
Species richness
0
20
40
60
80
100
120
Order richness
0
20
40
60
80
100
y=4.235x + 25.091
df=4915
R2=0.565
p<0.001

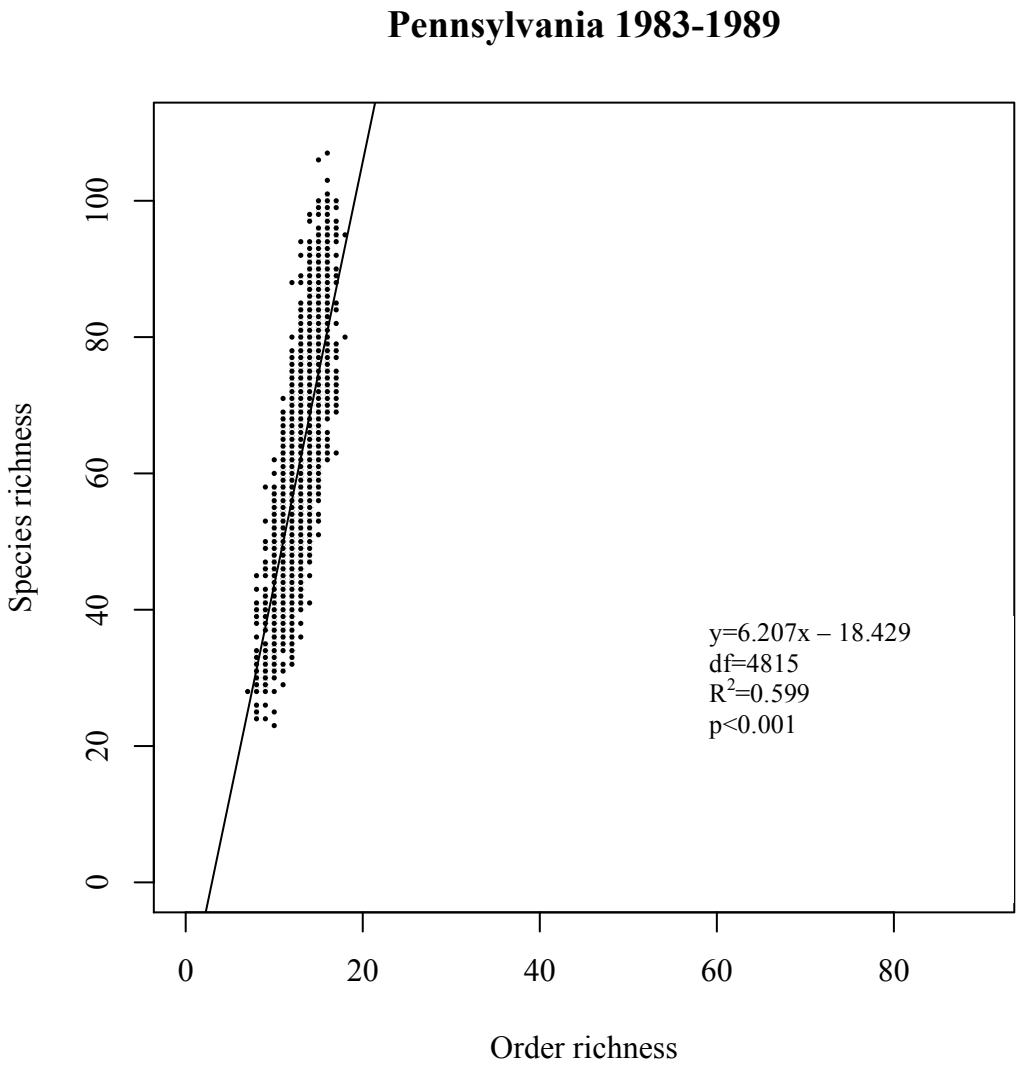
(a)


Pennsylvania 1983-1989
Species richness
Genus richness
0
20
40
60
80
100
y=1.269x – 5.273
df=4815
R2=0.974
p<0.001

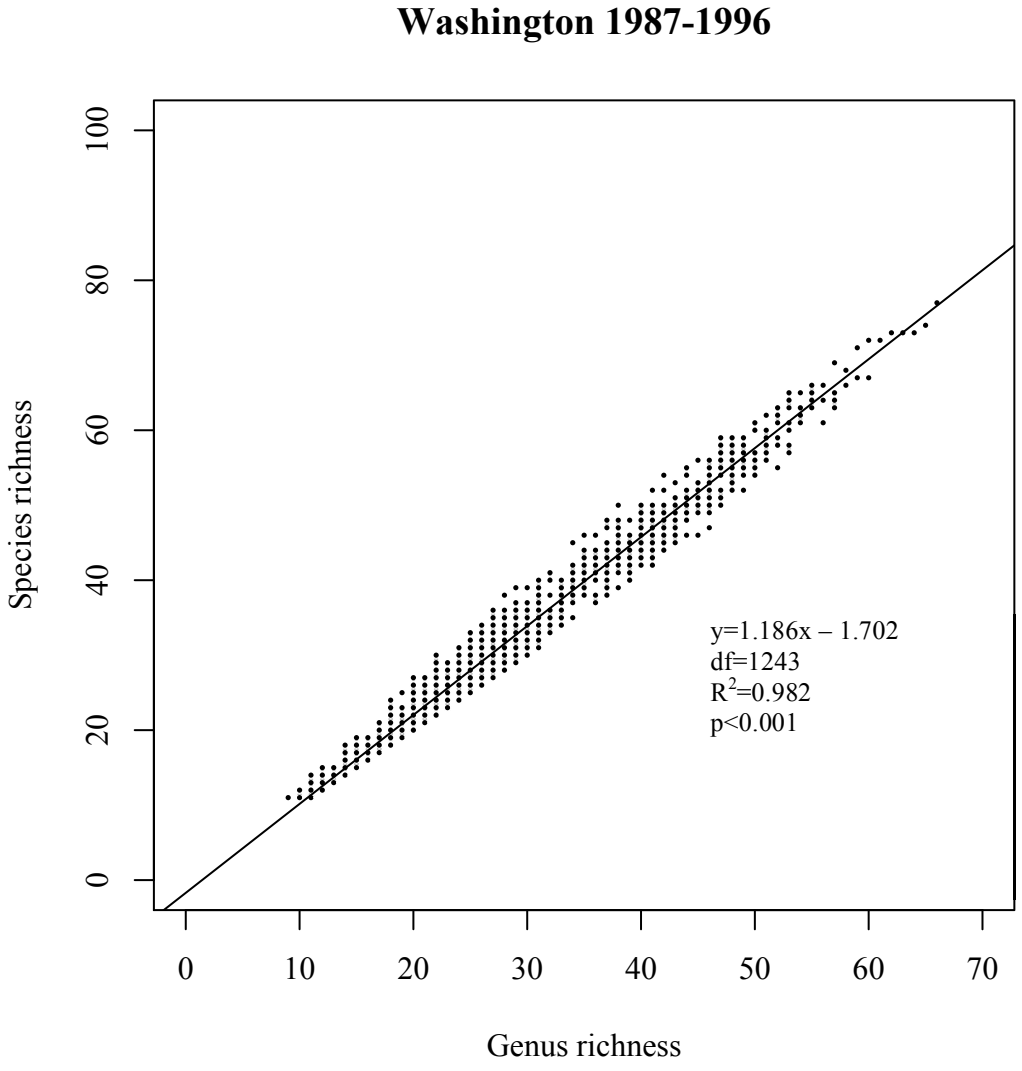
(b)


Pennsylvania 1983-1989
Species richness
Family richness
0
20
40
60
80
100
y=3.227x – 33.600
df=4815
R2=0.791
p<0.001

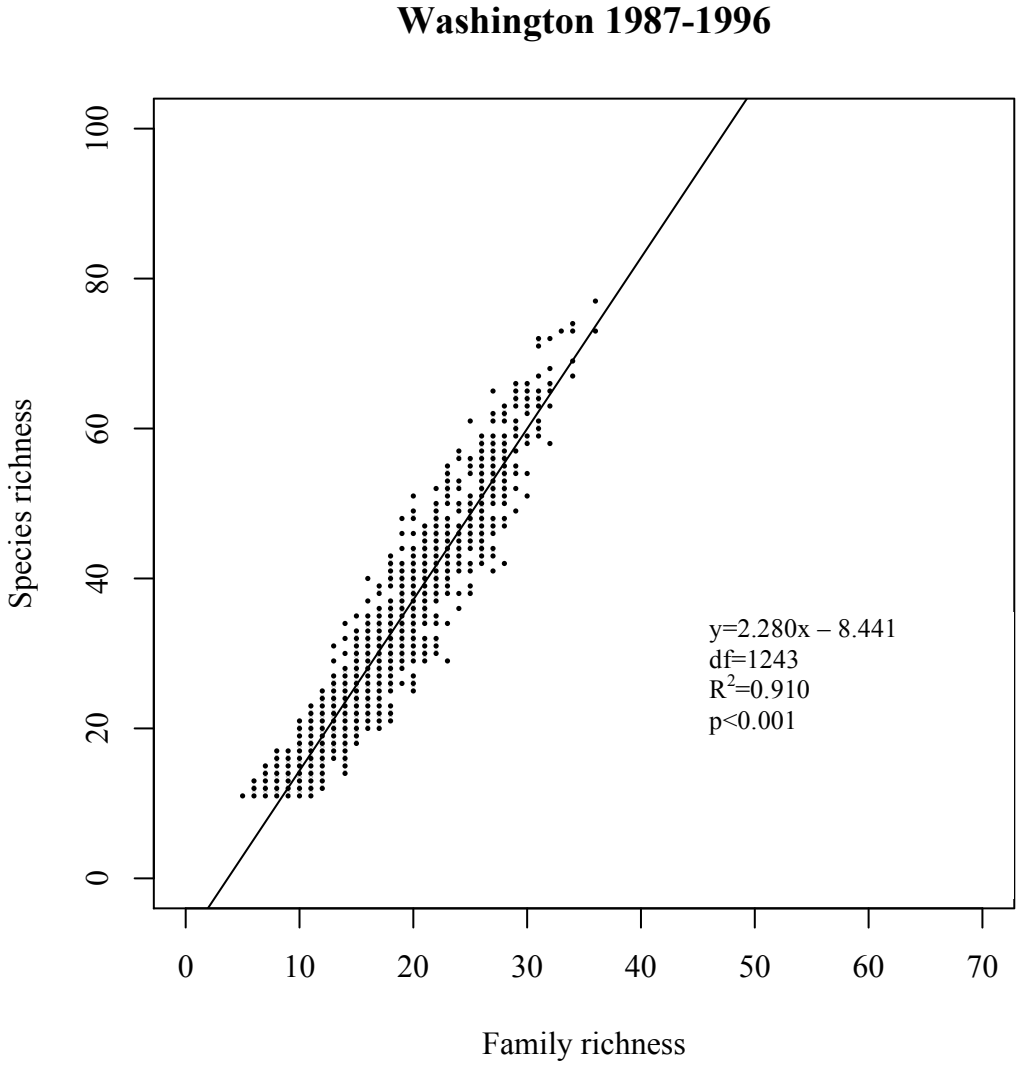
(c)


Pennsylvania 1983-1989
Species richness
0
20
40
60
80
100
Order richness
0
20
40
60
80
y=6.207x – 18.429
df=4815
R2=0.599
p<0.001

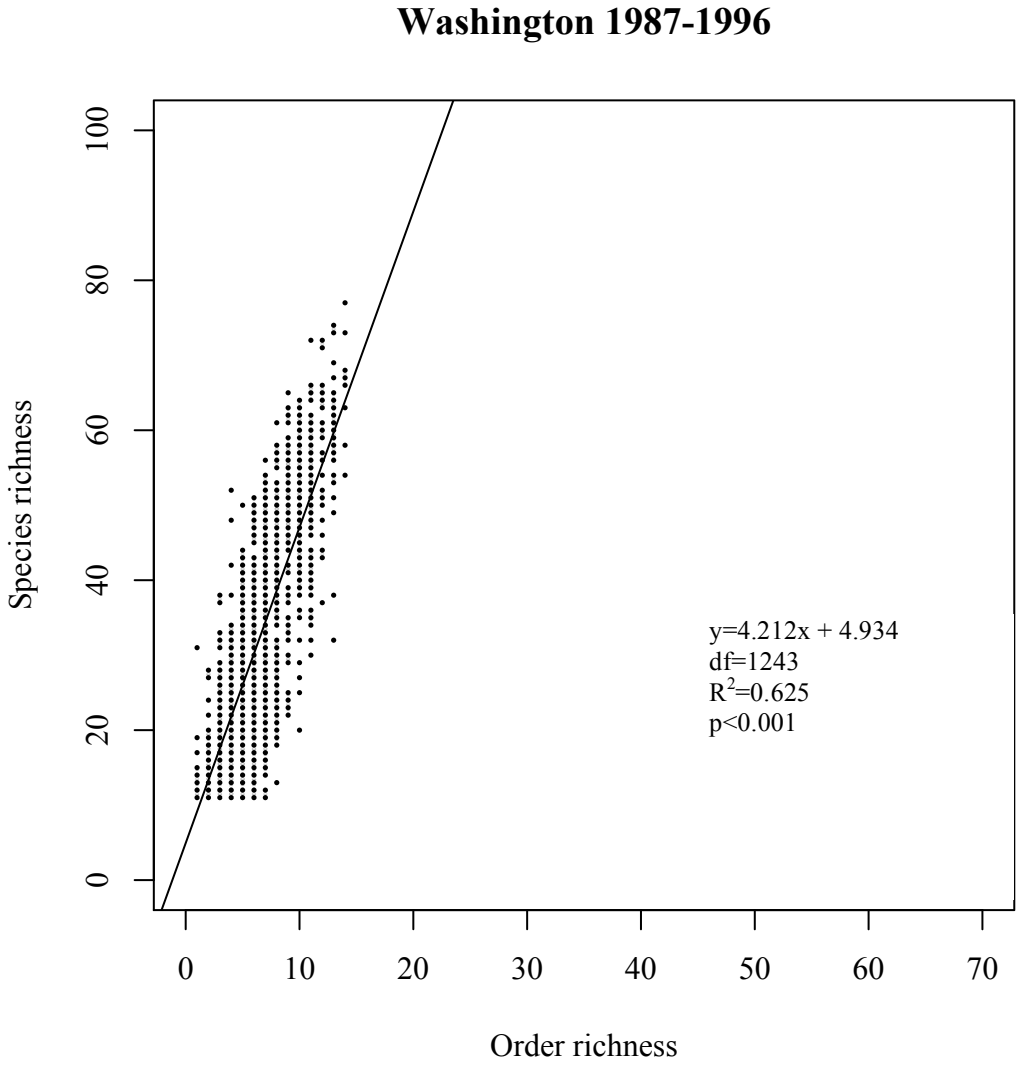
(a)


Washington 1987-1996
Species richness
Genus richness
0
20
40
60
80
100
0
10
20
30
40
50
60
70
y=1.186x – 1.702
df=1243
R2=0.982
p<0.001

(b)


Washington 1987-1996
Species richness
Family richness
0
20
40
60
80
100
0
10
20
30
40
50
60
70
y=2.280x – 8.441
df=1243
R2=0.910
p<0.001

(c)


Washington 1987-1996
Species richness
0
20
40
60
80
100
0
10
20
30
40
50
60
70
Order richness
y=4.212x + 4.934
df=1243
R2=0.625
p<0.001